



Army Science Conference Proceedings

12-15 June 1990

Volume III

Principal Authors N through Z

**This document has been approved for public
release and sale; its distribution is unlimited.**

19970929 065

FORM 9-74 (REV. 1-77) 1

Assistant Secretary of the Army
(Research, Development and Acquisition)
Department of the Army



DEPARTMENT OF THE ARMY
OFFICE OF THE ASSISTANT SECRETARY
WASHINGTON, DC 20310-0103



SARD-TR

30 JUL 1990

MEMORANDUM FOR SEE DISTRIBUTION

SUBJECT: Proceedings of the 1990 Army Science Conference

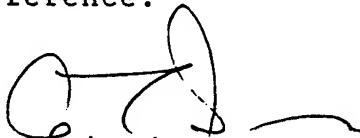
The 17th Army Science Conference was held at the Omni Durham Hotel and Convention Center, Durham, North Carolina, 12-15 June 1990. The conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army (DA). Additionally, it provided an opportunity for DA civilian and military scientists and engineers to present the results of their research and development efforts before a distinguished and critical audience.

These Proceedings of the the 1990 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted.

Our purpose for soliciting these papers was to:

- a. Stimulate the involvement of scientific and engineering talent within the Department of the Army;
- b. Demonstrate Army competence in research and development;
- c. Provide a forum wherein Army personnel can demonstrate the full scope and depth of their current projects; and
- d. Promote the interchange of ideas among members of the Army scientific and engineering community.

The information contained in these volumes will be of benefit to those who attended the conference and to others interested in Army research and development. It is requested that these Proceedings be placed in technical libraries where they will be available for reference.


George T. Singley III
Deputy Assistant Secretary
For Research and Technology

DISTRIBUTION:

Office of the Under Secretary of Defense for Research and Advanced Technology,
Wash, DC 20310-3080

Office of the Assistant Secretary of the Army (RD&A), ATTN: SARD-TR, Wash, DC
20310-0103

HQDA, Assistant Chief of Staff for Intelligence, ATTN: DAMI-ZA, Wash, DC
20310

HQDA, Deputy Chief of Staff for Logistics, ATTN: DALO-ZX, Wash, DC 20310-0500

HQDA, Deputy Chief of Staff for Operations & Plans, ATTN: DAMO-FD, Wash, DC
20310-0400

HQDA, Director of Information Systems for C4, ATTN: SAIS-ZA, Wash, DC
20310-0400

HQDA, Deputy Chief of Staff for Personnel, ATTN: DAPE-ZA, Wash, DC 20310-0300

Office of the Chief of Engineers, ATTN: CERD-M, 20 Massachusetts Avenue, N.W.,
Wash, DC 20314-1000

Office of the Surgeon General, ATTN: DASG-ZA, 5109 Leesburg Pike, Falls
Church, VA 22041-3258

COMMANDERS/DIRECTORS:

USA Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797

USA Operational Test & Evaluation Agency, 5600 Columbia Pike, Falls Church, VA
22041-5115

USA Information Systems Engineering Command, Ft. Huachuca, AZ 85613-5000

USA Strategic Defense Command, P. O. Box 15280, Arlington, VA 22215-0150

USA Materiel Command, ATTN: AMCLD, 5001 Eisenhower Avenue, Alexandria, VA
22333-0001

USA Materiel Systems Analysis Activity, ATTN: AMXSY-D, APG, MD 21005-5071

USA Armament, Munitions & Chemical Command, ATTN: AMSMC-CG, Rock Island, IL
61299-6000

USA Armament RD&E Center, ATTN: SMCAR-TD, Picatinny Arsenal, NJ 07806-5000

Fire Support Armament Cen, ATTN: SMCAR-FS, PA, NJ 07806-5000

Armament Engineering Dir, ATTN: SMCAR-AE, PA, NJ 07806-5000

Close Combat Armament Cen, ATTN: SMCAR-CC, PA, NJ 07806-5000

Chemical RD&E Center, ATTN: SMCCR-TD, APG, MD 21010-5423

USA Aviation Systems CMD, ATTN: AMSAV-GTD, 4300 Goodfellow Blvd, St. Louis, MO
63120-1798

USARTA, Ames Research Center, ATTN: SAVRT-D, Moffett Field, CA 94035-1099

USARTA, Aeroflightdynamics Dir, ATTN: SAVRT-AF, Moffett Field, CA 94035

USARTA, Aviation Applied Technology Dir, ATTN: SAVRT-TY, Ft Eustis, VA 23604

USARTA, Propulsion Dir, ATTN: SAVRT-PN, Lewis Research Center, 21000 Brook Park
Rd, Cleveland, OH 44135

USARTA, Aerostructures Dir, ATTN: SAVRT-SD, Hampton, VA 22665

Avionics R&D Activity, ATTN: SAVAA, Ft Monmouth, NJ 07703-5000

USA Communications-Electronics Cmd, ATTN: AMSEL-TDD, Ft Monmouth, NJ 07703-5001

Center for Command, Control & Communications System, ATTN: AMSEL-RD-C3,
Ft Monmouth, NJ 07703-5202

Center for Night Vision & Electro-Optics, ATTN: AMSEL-RD-NV-D, Ft Belvoir, VA
22060-5677

Electronics Warfare/RSTA Center, ATTN: AMSEL-RD-EW-D, Ft Monmouth, NJ
07703-5303

Signals Warfare Lab, ATTN: AMSEL-RD-SW-D, Vint Hill Farms Station, Warrenton,
VA 22186-5100

USA Laboratory Command, ATTN: AMSLC-TD, 2800 Powder Mill Road, Adelphi, MD 29783-1145
 Atmospheric Sciences Lab, ATTN: SLCAS-DD, WSMR, NM 88002-5501
 Ballistic Research Lab, ATTN: SLCBR-OD, APG, MD 21005
 Electronics Technology & Devices Lab, ATTN: SLCET-D, Ft Monmouth, NJ 07703-5000
 Electronic Warfare Vulnerability Assessment Lab, ATTN: SLCEW-M-D, WSMR, NM 8802-5513
 Harry Diamond Labs, ATTN: SLCHD-D, Adelphi, MD 20783-1197
 Human Engineering Lab, ATTN: SLCHE-H, APG, MD 21005-5001
 Materials Technology Lab, ATTN: SLCMT-D, Watertown, MA 02172-0001
 Army Research Office, ATTN: SLCRO-D, RTP, NC 27709-2211
 Army Research, Development & Standardization Group (Europe) Box 65, FPO New York 09510
 USA Missile Command, ATTN: AMSMI-R, Redstone Arsenal, AL 35898-5240
 USA Tank-Automotive Cmd, ATTN: AMSTA-CG, Warren, MI 48397-5000
 Aberdeen Proving Ground, ATTN: STEAP-OC, APG, MD 21005-5001
 Dugway Proving Ground, ATTN: STEDP-CO, Dugway, UT 84022-5000
 Electronic Proving Ground, ATTN: STEEP-DO, Ft Huachuca, AZ 85613-7110
 Aviation Development Test Activity, ATTN: STEBG-CO, Ft Rucker, AL 36362-5276
 Combat Systems Test Act, ATTN: STECS-CO, APG, MD 21005-5059
 White Sands Missile Range, ATTN: STEWS-CG, WSMR, NM 88002-5000
 USA Aviation Systems Cmd, ATTN: AMSAV-G, 4300 Goodfellow Blvd, St Louis, MO 63120-1798
 Belvoir RD&E Center, ATTN: STRBE-Z, Ft. Belvoir, VA 21005-5606
 Natick RD&E Center, ATTN: STRNC-Z, Natick, MA 01760-5000
 US Army Corps of Engineers
 Cold Regions Rsch & Eng Lab, ATTN: CRREL-TD, P. O. Box 282, Hanover, NH 03755-1290
 Construction Eng Rsch Lab, ATTN: CERL-Z, P. O. Box 4005, Champaign, IL 61820-135
 Engineer Topographic Labs, ATTN: ETL-TD, Ft Belvoir, VA 22060-5546
 Waterways Experiment Station, ATTN: WESTV, P. O. Box 631, Vicksburg, MS 39180-0631
 USA Medical R&D Command, ATTN: SGRD-ZA, Ft Detrick, Frederick, MD 21701-5012
 Aeromedical Rsch Lab, ATTN: SGRD-UAC, Ft Rucker, AL 36362-5292
 Inst of Dental Rsch, WRAMC, ATTN:SGRD-UDZ, Wash, DC 20307-5300
 Inst of Surgical Rsch, ATTN: SGRD-USZ, Ft Sam Houston, TX 78234-6200
 Letterman Army Inst of Rsch, ATTN: SGRD-ULZ, Presidio of San Francisco, CA 94129-6800
 Biomedical R&D Lab, ATTN: SGRD-UBZ, Frederick, MD 20701-5010
 Medical Rsch Inst of Chemical Defense, ATTN: SGRD-UV-ZA, APG, MD 21010-5425
 Medical Rsch Inst of Environmental Medicine, ATTN: SGRD-UEZ, Natick, MA 01760-5007
 Medical Rsch Inst of Infectious Diseases, ATTN: SGRD-UIZ-A, Ft Detrick, Frederick, MD 21701-5011
 Walter Reed Army Inst of Rsch, ATTN: SGRD-UWZ, Washington, DC 20307-5100
 USA Health Services Command, Ft. Sam Houston, TX 78234-61000
 USA Environmental Hygiene Agency, ATTN: HSHB-Z, APG, MD 21010-5422

USA Research Institute for the Behavioral and Social Sciences, ATTN: PERI-ZT,
5001 Eisenhower Ave, Alexandria, VA 22333-5600

ARI Field Unit, ATTN: PERI-IJ, POB 2086, Ft Benning, GA 31905

ARI Field Unit, ATTN: PERI-SB, POB 6057, Ft Bliss, TX 79906

ARI Field Unit, ATTN: PERI-SH, Ft Hood, TX 76544-5065

ARI Field Unit, ATTN: PERI-IK, Steele Hall, Ft Knox, KY 40121

ARI Field Unit, ATTN: PERI-SL, POB 290, Ft Leavenworth, KS 66027

ARI Field Unit, ATTN: PERI-IR, Ft Rucker, AL 36362-5000

ARI Field Unit, ATTN: PERI-IO, POB 5787, Presidio of Monterey, CA 93940

USA Training and Doctrine Command, ATTN: ATDO-ZT, Ft Monroe, VA 23651-5000

Combined Army Test Act, ATTN: ATCT-CG, Ft Hood, TX 76544-5065

COMMANDANTS:

USA Air Defense Cen, ATTN: ATZC-CG, Ft Bliss, TX 79916-5000

USA Armor Cen, ATTN: ATZK-CG, Ft Knox, KY 40121-5000

USA Aviation Cen, ATTN: ATZQ-CG, Ft Rucker, AL 36362-5000

USA Chemical Cen, ATTN: ATZN-CG, Ft McClellan, AL 36205-5000

USA Combined Arms Cen, ATTN: ATZL-SC, Ft Leavenworth, KS 66027-5000

USA Engineer Cen, ATTN: ATZA-CG, Ft Belvoir, VA 22060-5000

USA Infantry Cen, ATTN: ATZB-CG, Ft Benning, GA 31905-5000

USA Intelligence Cen, ATTN: ATSI-SA, Ft Hauchuca, AZ 85613-7000

USA Ordnance Cen, ATTN: ATSL-CMT, APG, MD 21005-5201

USA Signal Cen, ATTN: ATZH-CH, Ft Gordon, GA 30905-5000

USA Field Artillery Center, ATTN: ATZR-C, Ft Sill, OK 73503

SUPERINTENDENT:

US Military Academy, ATTN: Technical Library, West Point, NY 10996

COPIES FURNISHED:

Defense Advanced Research Projects Agency, 1400 Wilson Blvd, Arlington, VA
22209-2308

Defense Logistics Agency, Cameron Station, Alexandria, VA 22304-6183

Defense Technical Information Center, Cameron Station, Alexandria, VA
22304-6145

Lawrence Livermore National Lab, ATTN: L-191, POB 808, Livermore, CA 94550

Los Alamos National Lab, ATTN: Dir for Energy, Rsch & Tech, Los Alamos, NM
87545

NASA HQS, Suite 4237, 400 Maryland Ave, SW, Wash, DC 20546

National Science Foundation, 1800 G Street, NW, Wash, DC 20550

United Nations Library, ATTN: Acquisition Section, Room L-138A, New York, NY
10017

US NAVY

Naval Air Systems Command, Code AIR-03-D, 1411 Jefferson Davis Hwy,
Arlington, VA 22202-3000

Naval Rsch Lab, ATTN: Dir of Rsch, Washington, DC 20375

Office of Naval Rsch, Code 10, 800 North Quincy St, Arlington, VA 22217-5000

HQ US Marine Corps, Code RD-1, Washington, DC 29380-0001

US AIR FORCE

Air Force Systems Command, ATTN: Technical Director, Andrews AFB, Washington,
DC 20334-5000

Air Force Office of Scientific Research, ATTN: Tech Director, Bolling AFB,
Washington, DC 20332-6448

PROCEEDINGS
OF THE
1990 ARMY SCIENCE CONFERENCE

DURHAM, NORTH CAROLINA
12-15 JUNE 1990

VOLUME III
Principal Authors N through Z

TABLE OF CONTENTS

PROCEEDINGS OF THE 1990 ARMY SCIENCE CONFERENCE

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Abueme, Jeremias M.	See Grower, Marvin F.	II	103
Adams, George F.	Designer Energetic Materials	I	1
Adams, George F.	See Anderson, William R.	I	29
Adams, George F.	See Chabalowski, Cary F.	I	295
Aggarwal, Anita	See Sadoff, Jerald C.	III	153
Ahlrich, Randy C.	Resin Modified Pavement	I	13
Alexander, Millard H.	See Forch, Brad E.	I	573
Anderson, William R.	Chemistry of Solid Propellant Ignition	I	29
Anderson, William R.	See Sausa, Rosario C.	III	225
Anderton, Gary L.	See Ahlrich, Randy C.	I	13
Anitole, George	Development of Vehicle Woodland Camouflage Patterns Analogous to Background Contrast Highlights	I	45
Arfors, Karl E.	See Bruttig, Stephen P.	I	187
Armstrong, Lawrence E.	Exertional Heatstroke in Soldiers: An Analysis of Recovery Rates, Predisposing Factors and Residual Heat Intolerance	I	59
Ashley, P.	See Simonis, G. J.	III	287
Ashley, P. R.	See Bloemer, M. J.	I	99
Askew, E. Wayne	See Hoyt, Reed W.	II	193
Azzi, Kendra	See Crawford, Kenneth H.	I	389
Baer, Paul G.	See Coffee, Terence P.	I	323
Bancroft, William H.	An Inactivated Whole Virus Vaccine for the Prevention of Viral Hepatitis, Type A	I	73
Baron, Louis	See Sadoff, Jerald C	III	153

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Baskin, Steven I.	See Steinhaus, Ralph K.	III	409
Batten, Dee Ann	See O'Neill, Timothy	III	15
Beaudry, William T.	See Yang, Yu-Chu	III	547
Benton, John	Hierarchical Route Planner	I	87
Bernstein, Jeffrey S.	See Sausa, Rosario C.	III	209
Bickell, William H.	See Bruttig, Stephen P.	I	187
Binder, Michael	See Mammone, Robert J.	II	481
Binn, Leonard	See Bancroft, William H.	I	73
Black, E. D.	See Porter, William L.	III	93
Bloemer, Mark J.	Nonlinear Metal Composites for Guided Wave Devices in Optical Computing	I	99
Bloemer, Mark J.	See Bowden, Charles M.	I	113
Borgstrom, Per	See Bruttig, Stephen P.	I	187
Bowden, Charles M.	Nonlinear Optical Properties of Metallic Microparticle Composites	I	113
Bracuti, A. J.	Reaction Kinetic Model (PANDORA) with Reaction Induction Times for Molecular Collisions	I	129
Braddock, William D.	See Choi, K. K.	I	311
Braddock, William D.	See Smith, Doran D.	III	341
Bradley, James W.	See Murphy, Charles H.	II	593
Braitman, David J.	See Sparenborg, Steven	III	383
Braun, Christopher G.	MCT Solid-State Switching: Revolution in Power Conditioning	I	145
Brennecke, Lucas H.	See Sparenborg, Steven	III	383
Brewer, Ralph J.	See Perkins, Janet	III	55
Brink, Anne	See Benton, John R.	I	87

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Brodman, Bruce W.	See Gold, Kenneth	II	53
Brody, Philip S.	Laser Microscope using Phase-Conjugate Reconstruction from a Photorefractive Hologram	I	159
Bruchey, William J.	The Effect of Crystallographic Orientation on the Performance of Single Crystal Tungsten Sub-Scale Penetrators	I	171
Bruttig, Stephen P.	The Use of a Clinically Relevant Uncontrolled Hemorrhage Model to Study Massive Internal Bleeding and to Develop Appropriate Courses of Clinical Response to that Hemorrhage	I	187
Burrows, W. Dickinson	IV Fluidmaker: Preparation of Sterile Water for Injection in a Field Setting	I	201
Busciglio, Henry H.	The Potential of New Army Tests Improve Job Performance	I	213
Bushell, M.	Tactical Source Region Electromagnetic Pulse Simulation	I	227
Butler, Dwain K.	Ground Water Resource Assessments	I	237
Caldwell, John A. Jr.	The Impact of Atropine Sulfate on Flight Performance, Vision, Tracking, and Electroencephalographic Activity of Army Helicopter Pilots	I	253
Carignan, Yvon P.	The Morphology of Polymer Chains	I	269
Cartland, Harry E.	See Johnson, Thomas H.	II	239
Celmins, Aivars	Fuzzy Modeling of Shaped Charges Against Modern Armor	I	281
Chabalowski, Cary F.	Dynamical Behavior of Solid Explosives	I	295
Chabalowski, Cary F.	See Adams, George F.	I	1
Chandra, Suresh	See Daunt, Geraldine	I	451
Chang, Wayne H.	See Dornath-Mohr, Michelle A.	I	475
Cheng, Tu-Chen	See DeFrank, Joseph J.	I	461

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Chiang, Peter K.	See Smejkal, Ruthann M.	III	329
Chiu, D. S.	See Bracuti, A. J.	I	129
Choi, K. K.	GaAs Based Multiple Quantum Well 10 Micron Infrared Detectors	I	311
Christensen, Charles R.	See Bloemer, M. J.	I	99
Christensen, Charles R.	See Tanton, George A.	III	447
Chu, Yong-Kyu	See Schmaljohn, Connie S.	III	239
Clark, James H.	See Steinhaus, Ralph K.	III	409
Clark, Wm W. III	See Sharp, Edward J.	III	263
Coffee, Terence P.	Advances in Modeling Combustion Pro- cesses in Liquid Propellant Guns	I	323
Coffee, Terence P.	See Morrison, Walter F.	II	577
Cole, Melanie W.	Electrical and Microstructural Characterization of Ion Implanted Polycrystalline Silicon at High and Low Temperatures	I	339
Cole, Melanie W.	Microstructural Characterization of Semiconductor Materials as Related to Device Performance	I	349
Cole, Melanie W.	See Dornath-Mohr, Michelle A.	I	475
Combs, Roger J.	Theoretical and Experimental Aspects of a Short-Scan Interferometer for Remote Chemical Detection	I	363
Connolly, James J.	See Steeves, Diane M.	III	395
Cool, Terrill	See Sausa, Rosario C.	III	209
Cornell, John H.	A General Synthesis of Side Chain Liquid Crystalline Polymers	I	379
Coskunoglu, Beverly	See Schmidt, Wayne	III	253
Cotariu, Steven S.	See Kwan, Hon C.	II	333
Crane, Carl D.	See Sousk, Stephen F.	III	369

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Crawford, Kenneth	Program Support Environment for the Theater Construction Management System	I	389
Crawford, Robert M.	See Green, Shawn J.	II	81
Cross, Gerald	See Keniston, Richard C.	II	281
Cryz, Stanley	See Sadoff, Jerald C.	III	153
Cymerman, Allen	See Hoyt, Reed W.	II	193
Dagdigian, Paul J.	See Forch, Brad E.	I	573
Dagdigian, Paul J.	See Sausa, Rosario C.	III	225
Dahlstrom, Robert K.	Experimental Development of a Microwave Vlasov Mode Convertor	I	397
Dalrymple, Joel	See Schmaljohn, Connie S.	III	239
Danberg, James E.	Predicted Flight Performance of Base Bleed Projectiles	I	411
Dandekar, Dattatraya P.	Elastic Constants of a Graphite-Epoxy Composite by Ultrasonic Wave Velocity Measurements	I	427
Darrigrand, Andre A.	Reduction of Sweat Accumulation Rate with Pedal Antiperspirants	I	441
Daunt, Geraldine	Investigation of MIR Generation from 1 Micron Lasers Using Non-linear Techniques	I	451
Deberry, J.	See Garvin, Charles	II	27
DeCusatis, C.	See Kwan, H. C.	II	347
DeFatta, Richard P.	See Mullins, William D.	IV	65
DeFrank, Joseph J.	Hydrolysis of Toxic Organophosphorus Compounds by Enzymes from Halophilic Bacteria	I	461
Delgado, Ismael	See Keniston, Richard C.	II	281
DeLuca, Jane P.	See Armstrong, Lawrence E.	I	59
Dempsey, John K.	See Mullins, William D.	IV	65
Ditillo, John T.	See Combs, Roger J.	I	363

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Doherty, Tammy J.	See Bruttig, Stephen P.	I	187
Dominessy, Mary E.	See Lukas, Jeffrey H.	II	423
Dornath-Mohr, Michelle A.	Shallow, Low Temperature Au-Ge Ohmic Contacts to GaAs	I	475
Dubois, Doria R.	See Bancroft, William H.	I	73
Dulaney, Marland D., Jr.	Identification of a New Cyanide Countermeasure	I	491
Duncan, Fred	See Keniston, Richard C.	II	281
Dunne, C. P.	See Porter, William L.	III	93
Dutta, Mitra	Semiconductor Optical Waveguide Devices	I	507
Dutta, Mitra	See Choi, K. K.	I	311
Dutta, Mitra	See Cole, Melanie W.	I	349
Dutta, Mitra	See Smith, Doran D.	III	341
Eckart, Donald W.	See Dornath-Mohr, Michelle A.	I	475
Eckels, Kenneth	See Bancroft, William H.	I	73
Enriquez, John I., Sr.	See Keniston, Richard C.	II	281
Evans, Christopher	Metabolic Rescue with a Precursor for Adenylate Synthesis Stimulates Phosphate Utilization and Prolongs Survial in Hemorrhagic Shock	I	517
Falls, Terril C.	Automated Analysis of Interactions Between Terrain and Ground Force Mobility	I	533
Farrand, Timothy G.	See Magness, Lee S.	II	465
Fazi, Christian	Breakdown in Silicon p-n Junctions Due to Short Pulse, High-Field Electromagnetic Stress	I	543
Fedele, Paul D.	Aerosol Transport Modeling Indicates Hazards and Improved Protective Measures for Dusty Agents	IV	1
Ferry, Michael	See Daunt, Geraldine	I	451
Fleetwood, R.	See Bushell, M.	I	227

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Folkes, Patrick A.	Fluctuating Deep Level Trap Occupancy Mode for 1/f and Low-Frequency Noise in Semiconductor Transistors	I	559
Fong, Richard	New Lightweight, High Efficiency Explosively Formed Penetrator Warhead for Infantry Anti-Armor System	IV	17
Forch, Brad E.	Experimental and Computational Study of Collisions of Highly Excited Oxygen Atoms	I	573
Forch, Brad E.	Laser-Based Ignition of H ₂ /O ₂ and D ₂ /O ₂ Premixed Gases Through Resonant Multiphoton Excitation of H and D Atoms Near 243 nm: Initial Report of a Deuterium Isotope-Wavelength-Effect in Laser Ignition	I	587
Forch, Brad E.	Laser-Based Multiphoton Excitation Processes in Combustion Diagnostics	I	603
Forch, Brad E.	See Miziolek, Andrzej W.	II	555
Fortier, Anne H.	See Green, Shawn J.	II	81
Francesconi, Ralph	See Szlyk, Patricia C.	III	437
Franco, Raphael A.	A Very High Shock, Self-Contained Data Acquisition System	I	617
Franz, David R.	See MacDonald, Douglas A.	II	451
Friedl, Karl E.	Assessment of Body Weight Standards in Army Recruits	I	631
Fries, Joseph C.	Helicopter Rotor Blade Ballistic Damage Effects on Helicopter Dynamic Characteristics	I	645
Gale, John M.	See Perkins, Janet	III	55
Gales, Yolanda A.	See Smith, William J.	III	357
Gallman, Judith M.	The Validation and Application of a Rotor Acoustic Prediction Computer Program	II	1
Gallo, Benedict J.	Induction of Enhanced Strains of a Thermophile That Synthesize an OPA Anhydrase Effective in Hydrolyzing 4-Nitrophenyl Esters of Phenylphosphinate	II	17

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Gallo, Benedict J.	See Steeves, Diane M.	III	395
Garvin, Charles	Hybrid Approaches for Spread-Spectrum COMINT	II	27
Garvin, Charles	See Brody, Philip S.	I	159
Goff, John R.	Optical Associative Memory Based on Binary Phase Algebra and the Inner-Product Architecture	II	39
Gold, Kenneth	Biodegradation of Energetic Materials	II	53
Gordon, Daniel	See Sadoff, Jerald C.	III	153
Gordon, Richard K.	See Smejkal, Ruthann M.	III	329
Gowenlock, David A.	See Gallo, Benedict J.	II	17
Graham, Scott E.	Soldier Performance Research Project: Armor Field and SIMNET Tests	II	65
Grant, D. F.	See Porter, William L.	III	93
Graves, Bruce R.	See Mullins, William D.	IV	65
Green, Shawn J.	Identification of a Novel Cytokine-induced Effector Molecule Involved in the Intracellular Destruction of <u>Leishmania</u> , <u>Franciscella</u> and <u>Plasmodium</u> Sporozites: Nitric Oxide	II	81
Grillo, F. Gregory	Protection from Ricin Toxicity by Brefeldin A	II	95
Grisham, John A.	See Tanton, George A.	III	447
Grogl, Max	See Milhous, Wilbur K.	II	543
Gross, Clark L.	See Smith, William J.	III	357
Grower, Marvin F.	Development of Topically Applied Anti-Inflammatory Agents for Treating Acute Pulpal Inflammation in Combat Situations	II	103
Guidos, Bernard J.	Supercomputers Simulation of Supersonic Viscous Flow Over Pointed, Spherical, and Flat Tipped Shell	II	119
Gupta, N.	See Simonis, G. J.	III	287

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Haase, Randall R.	See Plowman, Kent M.	III	79
Hadwin, Larry J.	See Dahlstrom, Robert K.	I	397
Hales, Lyndell A.	See Marino, James N.	II	493
Halliday, John W.	Characterization of the Catalytic Site for G Agent Hydrolysis Using Electron Paramagnetic Resonance	II	135
Hamlet, Murray P.	See Darrigrand, Andre A.	I	441
Hansen, Richard L.	See La Sala, John E.	II	361
Haus, Joseph W.	See Bloemer, M. J.	I	99
Haus, Joseph W.	See Bowden, Charles M.	I	113
Hay, R.	See Simonis, G. J.	III	287
Hayes, Pamela G.	Vulnerability Analysis of Protective Structures Using Probabilistic Methods	II	149
Hergenroeder, Leo	See La Sala, John E.	II	361
Herczfeld, Peter R.	See Higgins, Thomas P.	II	163
Hewetson, John	See MacDonald Douglas A.	II	451
Higgins, Thomas P.	Optically Controlled Dielectric Re- sonator Oscillator for Millimeter Wave Applications	II	163
Hiller, Jack	Does OPTEMPO Increase Unit Readiness? An Objective Answer	II	171
Hisley, Dixie M.	See Opalka, Klaus O.	III	25
Hoffman, Steve L.	See Green, Shawn J.	II	81
Hoke, Charles H.	See Bancroft, William H.	I	73
Hoover, T. A.	A <u>Coxiella Burnetii</u> Repetitive DNA Element with Potential Diagnostic and Mutagenic Capabilities	II	181
Horwath, Edward J.	See Bruchey, William J.	I	171
Hoskin, Francis C.G.	See Steeves, Diane M.	III	395
Howe, Philip M.	Tank Firepower: An Analysis of Near- and Mid- Term Opportunities	IV	33

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Howard, Stephen L.	See Sausa, Rosario C.	III	209
Hoyt, Reed W.	Evaluating Combat Soldier Energy Balance in the Field with Stable Isotope and Ambulatory Monitoring Methods	II	193
Hsieh, Alex J.	Solvent Stress Cracking and Failure Mechanisms in Polyetherimide Composites	II	207
Hubbard, Roger W.	See Armstrong, Lawrence E.	I	59
Hubbard, Roger W.	See Szlyk, Patricia C.	III	437
Hudson, Thomas	See Grillo, F. Gregory	II	95
Ingram, James K.	See Franco, Raphael A.	I	617
Intaglietta, Marcos	See Bruttig, Stephen P.	I	187
Jackson, Ronald	See Darrigrand, Andre A.	I	441
Jaglowski, A. J.	See Singler, R. E.	III	315
Johnson, John L.	Object Extraction for Automatic Target Recognition	II	223
Johnson, John L.	See Anitole, George	I	45
Johnson, Robert A.	Laser-Induced Synthesis of Methoxy-methanol	II	253
Johnson, Ronald L.	See Anitole, George	I	45
Johnson, Thomas H.	Xenon Chloride Laser Scaling	II	239
Jones, Bruce H.	See Friedl, Karl E.	I	631
Jones, Kenneth A.	See Dornath-Mohr, Michelle A.	I	475
Jones, Tanya E.	See Hoyt, Reed W.	II	193
Jordan, Debbie B.	See Mullins, William D.	IV	65
Kalathil, Biju	See Crawford, Kenneth H.	I	389
Kaste, Richard C.	An Experimental Artillery Division Aid	IV	49
Kearl, Cyril E.	How Much Soldier Quality? Estimating Cost-Effective Recruit Selection Policy for a Smaller Army	II	265

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Keniston, Richard C.	The Sergeant Major Study Continued: Objective Health Risk Assessment By Laboratory Blood Tests	II	281
Khanna, Ravi	See Smith, Doran D.	III	347
Kim, Anderson	Sub-nanosecond Risetime High Power Photoconductive GaAs Switch and Its Transient Electric Field Profiles	II	307
Kim, Hie-Joon	Intrinsic Chemical Markers for Thermal Processing of Particulate Foods	II	295
Kingman, Priscilla W.	See Bruchey, William J.	I	171
Kirby, Stephen D.	See Steinhaus, Ralph K.	III	409
Kotlar, Anthony J.	See Anderson, William R.	I	29
Kotlar, Anthony J.	See Sausa, Rosario C.	III	209
Kotlar, Anthony J.	See Vanderhoff, John A.	III	491
Krasko, Genrich L.	The Effect of Hydrogen, Boron, Carbon, Phosphorus and Sulphur on Inter- granular Cohesion in Iron	II	321
Kroutil, Robert T.	See Combs, Roger J.	I	363
Kwan, H. C.	Experimental Polarization Dependence of Bragg Diffraction Using Surface Acoustic Waves in Lithium Niobate	II	347
Kwan, Hon C.	Measurement of the Refractive Indices of Fluoride Glasses Using Digital Refractometry	II	333
Kyle, Dennis E.	See Milhous, Wilbur K.	II	543
Lareau, Richard T.	See Dornath-Mohr, Michelle A.	I	475
La Sala, John E.	All-Optical Recombination Time Studies in Wide Gap Semiconductors	II	361
Leader, Haim	See Smejkal, Ruthann M.	III	329
Leavitt, R. P.	Novel Quantum-Well Geometries for Optoelectronic Device Applications	II	373

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Ledford, Mary E.	See Moore, Gerald L.	II	565
LeDuc, James W.	See Bancroft, William H.	I	73
Lehowicz, Lawrence G.	See Hiller, Jack H.	II	171
Leupold, Herbert A.	Novel Magnetic Field Sources for Micro, MM and Optical Wave Devices	II	389
Libelo, Louis F.	See Dahlstrom, Robert K.	I	397
Lieberman, Harris R.	See Hoyt, Reed W.	II	193
Little, J. W.	See Leavitt, R. P.	II	373
Litynski, D. M.	See Kwan, H. C.	II	347
Locke, Randy J.	See Miziolek, Andrzej W.	II	555
Locke, Randy J.	See Sausa, Rosario C.	III	209
Lombardi, Stephen J.	Nephila Clavipes Major Ampullate Gland Silk Proteins: Amino Acid Composition Analysis, Protein Sequencing, Construction and Se- quencing of Recombinant Genomic and cDNA Libraries	II	403
Losie, Lawrence D.	Calibration of Production Vulnera- bility/Lethality Models Based on Live Fire Test Results	II	409
Lovelette, Charles A.	See Cornell, John H.	I	379
Lukas, Jeffrey H.	Workload, Target Acquisition and Piloting Performance: Psychological and Physiological Predictors	II	423
Lunardini, V. J.	Measurement of Heat Loss From a Conduit-Type Heat Distribution System	II	435
MacDonald, Douglas A.	Analysis of Ventilatory Patterns Dur- ing Natural Recovery and Antibody- Medicated Reversal of Saxitoxin- Induced Apnea in the Guinea Pig	II	451
MacNeill, J.	See Porter, William L.	III	93
MacPherson, A. K.	See Bracuti, A. J.	I	129

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Mael, Fred A.	See White, Leonard A.	III	519
Magness, Lee S.	Deformation Behavior and Its Relationship to the Penetration Performance of High-Density KE Penetrator Materials	II	465
Malkin, Frank J.	See Lukas, Jeffrey H.	II	423
Mammone, Robert J.	Thin Polymers with Novel Dielectric Properties	II	481
Marchwicki, Ruth H.	See Bancroft, William H.	I	73
Marino, James N.	Research and Development by USACE to Enhance Logistics-Over-the Shore (LOTS) Operations	II	493
Martin, Anthony G.	See Dandekar, Dattatraya P.	I	427
Martin, Rodger K.	See Milhous, Wilbur K.	II	543
McFann, Howard H.	See Hiller, Jack H.	II	171
McGuire, D.	See Garvin, C.	II	27
McPherson, James C., Jr.	See Plowman, Kent M.	III	79
McPherson, James C., III	See Plowman, Kent M.	III	79
McQuaid, Michael J.	Laser Photochemical Studies of Intermediates in Energetic Materials Combustion	II	507
Mellouk, Sylvie	See Green, Shawn J.	II	81
Meltzer, Monte S.	See Green, Shawn J.	II	81
Merkel, George	See Bushell, M.	I	227
Mermagen, Wm. H. Sr.	See Murphy, Charles H.	II	593
Merrow, Clifton N.	See McQuaid, Michael J.	II	507
Michel, Rex R.	Tactical Estimate of the Situation: Past, Present, and Future	II	521
Middlebrook, John L.	Immunological Evaluation of Phospholipase A2 Neurotoxins	II	531

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Milhous, Wilbur K.	Circumventing Antimalarial Drug Resistance: Theoretical and Clinical Implications	II	543
Miller, Mary J.	See Sharp, Edward J.	III	263
Millnamow, Gregory A.	See Bruttig, Stephen P.	I	187
Miziolek, Andrezej W.	Novel Laser-Based Detector for Gas Chromatography	II	555
Miziolek, Andrzej W.	See Forch, Brad E.	I	573
Miziolek, Andrzej W.	See Forch, Brad E.	I	587
Miziolek, Andrzej W.	See Forch, Brad E.	I	603
Miziolek, Andrzej W.	See McQuaid, Michael J.	II	507
Miziolek, Andrzej W.	See Sausa, Rosario C.	III	209
Miziolek, Andrzej W.	See Sausa, Rosario C.	III	225
Monty, Richard A.	See Lukas, Jeffrey H.	II	423
Moore, Gerry L.	Additive Solutions for the 21 Day Preservation of Previously Frozen Red Blood Cells	II	565
Morris, Jeffrey B.	See Forch, Brad E.	I	603
Morris, Jeffrey B.	See Miziolek, Andrzej W.	II	555
Morrison, Walter F.	A Modified Lagrange Pressure Gradient for the Regenerative Liquid Propellant Gun	II	577
Morrison, Walter F.	See Coffee, Terence P.	I	323
Mullins, William D.	Test and Evaluation of Man Portable Lasers for Infantry Use IV	IV	65
Murphy, Charles H.	Side Moment Exerted by a Spinning, Coning, Highly Viscous Liquid Payload	II	593
Nacy, Carol A.	See Green, Shawn J.	II	81
Nattress, D.	See Porter, William L.	III	93
Nelson, Douglas C.	See Burrows, W. Dickinson	I	201
Nelson, Douglas C.	See Fedele, Paul D.	IV	1
Nelson, James H.	See Burrows, W. Dickinson	I	201

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Neubert, Christopher J.	See Anitole, George	I	45
Newman, Peter G.	See Choi, K. K.	I	311
Newman, Peter G.	See Cole, Melanie W.	I	349
Ng, William	See Fong, Richard	IV	17
Nietubicz, Charles J.	See Danberg, James E.	I	411
Nietubicz, Charles J.	See Sahu, Jubaraj	III	165
Nord, Roy D.	See Kearl, Cyril E.	II	265
Nord, Roy D.	See White, Leonard A.	III	519
Nusca, Michael J.	Computational Fluid Dynamics Application to the Aerodynamics of Symmetric Sabot Discard	III	1
O'Benar, John D.	See Bruttig, Stephen P.	I	187
O'Neill, Timothy R.	Predicting Target Detection Skill	III	15
Oatman, Lynn C.	See Lukas, Jeffrey H.	II	423
Oduola, Ayoade M. J.	See Milhous, Wilbur K.	II	543
Opalka, Klaus O.	Optical Studies of the Flow Start-Up in Convergent/Divergent Nozzles	III	25
Paolella, Arthur	See Higgins, Thomas P.	II	163
Pastore, Robert	See Braun, Christopher G.	I	145
Paustian, Paul W.	See Plowman, Kent M.	III	79
Pellicore, Linda S.	See Dulaney, Marland D., Jr.	I	491
Pergantis, Charles G.	See Perkins, Janet	III	55
Perkins, Janet S.	Directed Energy Warfare: Advances in Protective Armor Materials	III	55
Pignatiello, J. J.	See Porter, William L.	III	93
Pin, Francois G.	See Sousk, Stephen F.	III	369
Pleban, Robert J.	Ranger and Special Forces Research: Implications for Low Intensity Conflict Training and Operations	III	69

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Plowman, Kent M.	Acutely Enhanced Burn Recovery by Post-Event Non-Ionic Surfactant Treatment in a Rat Model	III	79
Porter, William L.	Role of Ascorbic Acid and Tin in Maillard Browning of Wet-Pack Fruits	III	93
Purchase, K.	See Simonis, G. J.	III	287
Rajan, K. S.	See Halliday, John W.	II	135
Reynolds, Katy	See Darrigrand, Andre A.	I	441
Rice, Betsy M.	See Adams, George F.	I	1
Rice, Betsy M.	See Chabalowski, Cary F.	I	295
Rice, Betsy M.	See Trevino, S. F.	III	461
Richardson, Judith D.	A Layered Architecture for Interfac- ing Ada and SQL	III	109
Richmond, Paul W., III	Vehicle Motion Resistance Due to Snow	III	125
Rinker, Jack N.	Hyperspectral Imagery - A New Technique for Targeting and Intelligence	III	137
Roberts, Donald E.	See Darrigrand, Andre A.	I	441
Rohrbaugh, Dennis K.	See Yang, Yu-Chu	III	547
Rosano, D.	See Porter, William L.	III	93
Rossan, Richard N.	See Milhous, Wilbur K.	II	543
Runner, Royce R.	See Plowman, Kent M.	III	79
Ruth, Brian G.	See Dahlstrom, Robert K.	I	397
Sabol, Mark A.	See Wisher, Robert A.	III	533
Sadoff, Gerald C.	Development of Vaccines Against Malaria	III	153
Sahu, Jubaraj	Three Dimensional Flow Calculations for a Projectile with Standard and Dome Bases	III	165

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Sanders, Kenneth M.	See Smith, William J.	III	357
Santiago, Joseph M.	Finite Element Analysis of the Dynamic Thermal Buckling of a Thin Wall Cylinder	III	179
Sass, David T.	See Shepard, Steven M.	III	279
Satterwhite, Melvin B.	Spectral Luminescence of Camouflage Fabrics and Paints	III	195
Sausa, Rosario C.	Excimer Laser Photodissociation of Selected Oximes: Search for H ₂ CN Radical	III	225
Sausa, Rosario C.	Laser Spectroscopic and Mass Spectrometric Studies Propellant-Like Low Pressure Flames	III	209
Sausa, Rosario C.	See McQuaid, Michael J.	II	507
Schmaljohn, Connie	Expression of Hantaan Viral Antigens for Vaccine Development	III	239
Schmidt, Wayne	KNOWLEDGE WORKER SYSTEM An Investment for the Army's Future	III	253
Schuster, Brian G.	See Milhous, Wilbur K.	II	543
Scotland, Paula M.	See Gallo, Benedict J.	II	17
Seng, George F.	See Grower, Marvin F.	II	103
Sennett, M. S.	See Singler, R. E.	III	315
Sharp, Edward J.	Image Processing Applications Via Photorefraction	III	263
Shepard, Steven M.	Thermal Imaging at Above Frame Rate Frequencies	III	279
Shockley, Deborah K.	See Goff, John R.	II	39
Sils, Ingrid V.	See Szlyk, Patricia C.	III	437
Silva, Jay	See Busciglio, Henry H.	I	213
Simonis, G. J.	Optoelectronic Generation, Control, and Distribution of Microwaves	III	287

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Sims, S. Richard F.	An Integrated Target Acquisition System for Fire Control and Autonomous Acquisition	III	303
Singler, R. E.	Poly(organophosphazenes) for Non-linear Optical Applications	III	315
Sjogren, Maria H.	See Bancroft, William H.	I	73
Smejkal, Ruthann M.	Binary Antidotes for Organophosphate Chemical Warfare Agents	III	329
Smith, Doran D.	The Effects of Electron Beam Processing on High Electron Mobility Transistors	III	341
Smith, Doran D.	A Selectively-Contacted Dual Channel High Electron Mobility Transistor	III	347
Smith, M.	See Bushell, M.	I	227
Smith, William J.	Sulfur Mustard-induced Biochemical Alterations in Proliferating Human Cells in Culture	III	357
Sousk, Stephen F.	Motion Planning for the Universal Self-Deployable Cargo Handler	III	369
Sparenborg, Steven	The Excitatory Amino Acid Antagonist MK-801 Prevents Nerve Agent-induced Neuropathology Even When Given After the Onset of Convulsions	III	383
Stanley, Ann E.	See Johnson, Robert A.	II	253
Steeves, Diane M.	Stereospecificity of Microbial Enzymes for G-Agent Detoxification	III	395
Steinhaus, Ralph K.	Formation of Methemoglobin and Metmyoglobin Using 8-Amino-quinoline Derivatives or Sodium Nitrite and Subsequent Reaction with Cyanide	III	409
Stensby, John	See Tanton, George A.	III	447
Still, G.	See Bushell, M.	I	227
Sturek, Walter B.	See Weinacht, Paul	III	503

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Swinson, Mark L.	Military Robotics: A Technology Assessment From a User's Perspective	III	423
Szafraniec, Linda L.	See Yang, Yu-Chu	III	547
Szlyk, Patricia C.	An Innovative Fluid Delivery System for Chemical Protective Clothing: An Evaluation	III	437
Tanton, George A.	UV-IR Detector and Focal Plan Array Material Evaluation Using Faraday Rotation	III	447
Taub, Irvin A.	See Kim, Hie-Joon	II	295
Thompson, Donald L.	See Chabalowski, Cary F.	I	295
Thompson, Thomas J.	See Pleban, Robert J.	III	69
Trevino, S. F.	A Determination of the Inter-molecular Potential of Nitromethane	III	461
Underwood, Robert B.	See Falls, Terril C.	I	533
Valdes, James J.	Direction of Toxins with a Reversible Biosensor	III	477
Valentine, Patrick J.	See Pleban, Robert J.	III	69
Vanderhoff, John A.	Simultaneous Determination of Temperatures of OH Concentrations in a solid Propellant Flame	III	491
Vlahocos, Constaine	See Dulaney, Marland D., Jr.	I	491
Vodkin, M. H.	See Hoover, Timothy A.	II	181
Vogel, James A.	See Friedl, Karl E.	I	631
Wade, Charles E.	See Bruttig, Stephen P.	I	187
Walker, Clinton	See Busciglio, Henry H.	I	213
Walker, John E.	See Halliday, John W.	II	135
Walker, John E.	See Steeves, Diane M.	III	395
Walkinshaw, John W.	See Perkins, Janet	III	55

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Ward, Alford L.	See Fazi, Christian	I	543
Weinacht, Paul	Navier-Stokes Predictions of Pitch Damping for High L/D Finned Kinetic Energy Projectiles	III	503
Weinacht, Paul	See Guidos, Bernard J.	II	119
Weiner, M.	See Kim, Anderson	II	307
Weiner, Maurice	See Braun, Christopher G.	I	145
Westgate, Roger C.	See Fazi, Christian	I	543
Wheeler, Thomas J.	See Richardson, Judith D.	III	109
White, Leonard A.	Setting Enlistment Standards on the ABLE to Reduce Attrition	III	519
Wiesmann, William	See Evans, Christopher P.	I	517
Williams, Jim C.	See Hoover, Timothy A.	II	181
Willingham, Reginald A.	See Singler, R. E.	III	315
Wisher, Robert A.	Predicting the Decay of Mobile Subscriber Equipment (MSE) Operator Skills	III	533
Wisler, John A.	See Dulaney, Marland D., Jr.	I	491
Wisniewski, Henry L.	See Santiago, Joseph M.	III	179
Wolfe, A. David	See Smejkal, Ruthann M.	III	329
Wood, Gary L.	See Sharp, Edward J.	III	263
Woontner, Susan	See O'Neill, Timothy R.	III	15
Wren, Gloria P.	See Coffee, Terence P.	I	323
Yang, Yu-Chu	Oxidative Decontamination of the Chemical Nerve Agent	III	547
Yip, Pearl W.	See Perkins, Janet	III	55
Youmans, R.	See Kim, Anderson	II	307
Younger, M.	See Garvin, Charles	II	27
Zeto, R.	See Kim, Anderson	II	307

Computational Fluid Dynamics Application to the Aerodynamics of Symmetric Sabot Discard

* Michael J. Nusca

US Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

I. INTRODUCTION

Currently, the most widely utilized design for kinetic energy, antitank applications is the gun launched, fin-stabilized, long-rod projectile. The cross-sectional diameter of the rod is smaller than the diameter of the gun bore. Fins span the area between the rod and the gun tube. Thus a sabot is required to reduce in-bore balloting of the projectile. Once free of the gun tube, the sabot must be discarded in order to permit unconstrained, low-drag flight to the target, see Figure 1. Typically, the sabot is divided into three or four components along axial planes. For smooth bore gun tubes, these components separate from the projectile under the action of elastic and aerodynamic loads. During the separation, both mechanical interference and gasdynamic forces may result in alteration of the projectile's trajectory. This alteration can result in unacceptable loss of accuracy at the target.

Schmidt and Shear^{1,2} have demonstrated that aerodynamic interference, generated by the sabot petals, can be a significant source of projectile launch disturbance. Perturbations to the projectile's trajectory are magnified by geometric asymmetry in the discard pattern and by extended periods during the flight when the sabot components are in close proximity to the projectile. A detailed understanding of the three-dimensional shock/boundary-layer interference flowfield between the sabot and the projectile is not available.

Schmidt and Plostins^{3,4} have conducted an extensive experimental program to investigate the aerodynamics of sabot discard. During these tests a representative projectile/sabot model was sting-mounted in the NASA Langley Unitary Plan wind tunnel facility. The model configuration included a cone-cylinder projectile (without fins) at zero yaw and zero axial spin, with three sabot components of simplified geometry.

Figure 2 shows a schematic of the wind tunnel model and associated nomenclature. Surface pressures were recorded on the projectile and sabot petals (both inner and outer surfaces) for a freestream Mach number (ratio of velocity to speed of sound) of 4.5 and Reynolds number (ratio of inertial to viscous forces) of 6.6 million. A typical flight Reynolds number of about 89 million could not be reproduced in the tunnel. Figure 3 is a schematic of a portion of the test series showing sabot petals symmetrically located at four positions near the projectile and at angle of attack. This "simulated sabot discard" sequence provides an good basis for comparison of computed and measured surface pressures during discard. Comparison of test data with simple Newtonian flow theory demonstrated that the effects of shock/boundary-layer interaction must be included in order to adequately reproduce measurements.

This paper describes computational fluid dynamics (CFD) solutions of the three-dimensional Navier-Stokes equations applied to the aerodynamics of symmetric sabot discard. For symmetric discard, multiple sabot petals are assumed to follow a similar trajectory away from the projectile, and the projectile is assumed to be at zero yaw. Figure 4 is an illustration of this configuration for three sabot petals. The computational domain for symmetric discard can be limited to a smaller portion of the flowfield around the configuration. Thus requirements for three-dimensional simulations such as computational grid size, computer memory, and computer run time are reduced. For asymmetric discard the computational domain would be greatly expanded with a corresponding increase in computer requirements. The portion of the launch cycle which involves strong aerodynamic interference between the projectile and the sabot components is examined. Thus simulations are performed for small vertical separation of the sabot petals from the projectile and sabot angle of attack $\leq 18^\circ$. Numerical simulations reported here were performed using the wind tunnel model sabot configuration illustrated in Figures 2 and 3. Zero axial spin is assumed for the projectile/sabot configuration.

II. COMPUTATIONAL APPROACH

The CFD approach described in this section can be used to predict the compressible flowfield around single and multiple component non-axisymmetric projectiles by solving the Reynolds-averaged 3D Navier-Stokes equations. The CFD code was developed by Chakravarthy.⁵ The Navier-Stokes equations are written using the perfect gas assumption, however, versions of the code for a real gas (e.g. equilibrium air or non-equilibrium gas models) have been documented as well.^{6,7} Both laminar and turbulent flows can be investigated, thus an adequate turbulence model⁸ is required for closure. In addition, backflow regions can be present, thus a backflow turbulence model⁹ is included. The equations are discretized using finite volume approximations.

The code uses a class of numerical algorithms termed total variational diminishing, or TVD, which do not require the inclusion of explicit smoothing or global dissipation functions to achieve numerical stability. The resulting set of equations (see Reference 10 for two-dimensional versions) is solved using an implicit, factored, time-stepping algorithm. The solution takes place on a computational grid that is generated around the projectile and sabot petals in zones, where the zonal boundaries can be made transparent to the flowfield calculation. This code has been previously employed in the solution of subsonic, transonic, supersonic and mixed flow problems including complex three-dimensional and multi-body configurations by Chakravarthy et. al.^{5,10-14}

1. Computational Algorithm

The spacial discretization technique for the equations of motion must be reliable and robust in order to successfully capture the complex physics of projectile/sabot interacting flowfields. The TVD formulation for the convection terms (the hyperbolic part of the time-dependent Navier-Stokes equations), along with a special treatment of the diffusion terms, provides an appropriate simulation. In recent years, TVD formulations have been constructed for shock-capturing finite-difference methods.⁵ Near large gradients in the solution (extrema), TVD schemes automatically reduce to first-order accurate discretizations locally, while away from extrema they can be constructed to be of higher-order accuracy. This local effect, which is necessary to prevent the total variation from increasing, restricts the maximum global accuracy possible for TVD schemes to third order for steady-state solutions. These methods manifest many properties desirable in numerical solution procedures. By design, they avoid numerical oscillations and "expansion shocks" while at the same time being higher-order (more than first-order) accurate. TVD formulations are also based on the principle of discrete or numerical conservation, which is the numerical analog of physical conservation of mass, momentum, and energy. Thus, TVD schemes can "capture" discontinuities with ease and high resolution. At a fundamental level, they are based on upwind schemes; therefore, they closely simulate the signal propagation properties of hyperbolic equations. Schemes based on the TVD formulation are completely defined. In contrast, central difference schemes involve global dissipation terms for stability and have one or more explicit coefficients that must be judiciously chosen to achieve desirable results.

2. Computational Grid

Numerical simulation of the interacting flowfield about projectile/sabot combinations is complicated by the non-axisymmetric geometry (Figure 4). For even simplified sabot configurations, sharp corners severely hamper conventional grid generation

schemes that require one set of grid lines to be tangential and another set to be normal to the surface. Projectile/sabot geometries are more easily gridded using the zonal approach. The configuration is divided into zones of simple geometric shape. In each zone an algebraic grid is generated with grid clustering near the surface and high flow gradient regions. In the zonal approach, the computational method and computer program are constructed in such a manner that each zone may be considered as an independent module, interacting with other zones before or after the information corresponding to each zone is updated one cycle. In addition, the zonal boundaries can be made transparent to flowfield phenomena (e.g. shock waves).

A typical 6-zone grid used for computations described in this paper is designed as follows: grid zone 1 covers the projectile from nose to base; zone 2 covers the area between zone 1 and the inner surface of the sabot petal; zone 4 covers the area between the outer surface of the sabot petal and the uppermost extent of the computational space; zones 5 and 6 cover the projectile and sabot base regions, respectively. Zones 1 thru 6, excluding zone 3, extend from 0 to 60° in the azimuthal direction. Grid zone 3 covers the area between the azimuthal face of the sabot petal and the end of the computational area (see Figure 4). The entire 6-zone grid consists of 320,000 points and requires 11 million word of memory on the BRL CRAY-2 supercomputer. Grid clustering was used along all wall boundaries to resolve boundary layer profiles and near geometric discontinuities to resolve flow gradients. The same grid can be used for cases where the sabot petal is pitched to an angle of $\leq 18^\circ$, due to the use of a no-reflection boundary condition on the uppermost surface of the computational space. These grid dimensions are assumed to be sufficient for preliminary computational simulations where relatively short computer run time was required. Grid refinement studies are underway.

III. RESULTS

This paper presents numerical simulations for the wind tunnel projectile/sabot model. The freestream Mach number and Reynolds Number (Re) are 4.5 and 6.6 million, respectively. Calculations are presented for laminar (i.e. viscosity depends on temperature only) and inviscid (i.e. $Re = \infty$) flow modeling since the current grid spacings are inadequate for turbulent flow calculations. Converged solutions required about 6 CPU hours for inviscid flow and 18 CPU hours for laminar flow modeling on the BRL CRAY-2 computer.

Figure 5 shows pressure contours for the forward part of the projectile/sabot configuration, in the pitch plane (azimuthal angle, $\phi = 0$). In this case the base of the sabots are aligned with the projectile base, $x/D = 0$, the sabots and projectile are ver-

tically separated by $y/D = .75$, and the sabot petals are at zero yaw. The maximum diameter of the projectile is denoted D . Horizontal lines running from $x/D = 0 - 5.62$ are zonal boundaries. The presence of large flow gradients (e.g. shock waves) are indicated by the clustering of contour lines. An oblique shock emanates from the nosetip of the projectile while a normal shock is slightly detached from the leading edge of the sabot petal. These shocks are followed downstream by a complex interaction and reflection pattern. A flow expansion at the cone-cylinder junction is followed by a strong oblique shock on the projectile. Thus a high pressure, low speed and perhaps recirculating flow region may be present on the projectile beneath the sabot petals. This region terminates with a second normal shock that appears at the end of the beveled section of the sabot, $x/D = 4.22$. This normal shock intersects the projectile as well.

Figure 6 shows the sonic (Mach = 1) Mach contours for the same configuration. Horizontal lines running from $x/D = 0 - 5.62$ are zonal boundaries. These contours indicate that a region of subsonic (Mach < 1) flow extends from $x/D = 2.8 - 4.22$, between the sabot petals and the projectile.

Figure 7 shows the pressure distribution over the projectile and sabot surfaces in the pitch plane from computations and experimental measurements.³ These laminar flow simulations for the projectile surface agree favorably with the magnitude and location of a measured pressure peak ($x/D = 4.22$) and an elevated pressure (low speed flow) region between the flow expansion at the cone-cylinder junction and the peak ($2 \leq x/D \leq 4.22$). The location of this pressure peak corresponds to the termination of a low speed flow region on the projectile (see Figure 6). Downstream of the pressure peak, the agreement between computation and measurement is also favorable. On the inner surface of the sabot petals numerical simulation adequately predicts the pressure levels and trends on the sabot slant surface ($2.75 \leq x/D \leq 3.94$). Pressure levels on the rest of the sabot section agree with measurements including a pressure rise at $x/D = 5.5$. The computed magnitude of this rise is somewhat larger than observed.

Figure 8 shows the pressure distribution over the projectile and sabot surfaces in the pitch plane from computations and experiment. In this case the sabots and projectile are vertically separated by $y/D = .5$, and the sabot petals are at 4° of yaw. The pressure levels and trends are similar to those of Figure 7 but with slightly elevated values from $x/D = 7 - 11$ where the bodies are in closer proximity. The magnitude of the pressure peak at $x/D = 4.22$ on the projectile is smaller for this case since only one of the three sabot petals was mounted, while flow splitter plates were attached to the projectile along symmetry planes ($\phi = \pm 60^\circ$). The computation did not include these splitter plates.

Figure 9 shows the pressure distribution over the projectile and sabot surfaces in the pitch plane from inviscid flow computations and experimental measurements. In-

viscid flow simulations require significantly less computer time by excluding the viscous terms in the Navier-Stokes equations. These simulations predict a smaller magnitude pressure peak and lower pressure forward of the peak, indicating that these are perhaps viscous phenomena. Inviscid simulations predict a much larger sabot leading edge pressure and lower pressure on the sabot slant surface ($2.75 \leq x/D \leq 3.94$). Thus the inviscid computation is not an adequate substitute for the fully viscous simulation.

Calculations for an axisymmetric configuration are presented in Figures 10 and 11. The axisymmetric configuration is generated by joining the three sabot petals into a tube that is concentric with the projectile. The axisymmetric configuration requires only a two-dimensional simulation which reduced the required computer time from 18 hours to 3 hours on the BRL CRAY-2. Figure 10 shows that the axisymmetric geometry generates a normal shock that is significantly detached from the sabot (tube). As a result, this shock interferes with the flowfield at the cone-cylinder junction on the projectile. A strong oblique shock is generated at this location. This flowfield is significantly different from that for the actual geometry (see Figure 5).

Figure 11 shows the pressure distribution over the projectile and sabot surfaces in the pitch plane for the axisymmetric configuration. Several pressure peaks are generated on the projectile surface. One of these peaks is similar in location to a measured pressure peak. The pressure measurements on the sabot slant surface are reproduced by the calculation, however pressures between the sabot (tube) and the projectile are much higher than measurements. These results indicate that the axisymmetric simulation predicts a choked flow between the projectile and the sabot and is not an adequate substitute for the non-axisymmetric simulation.

IV. CONCLUSIONS AND FUTURE WORK

Computational fluid dynamics (CFD) solutions of the three-dimensional Navier-Stokes equations have been applied to the aerodynamics of symmetric sabot discard. The portion of the launch cycle which involves strong aerodynamic interference between the projectile and the sabot components was examined. A model configuration was considered that included a cone-cylinder projectile at zero yaw with three sabot components of simplified geometry. Computed and measured projectile and sabot surface pressures are in good agreement for a freestream Mach number of 4.5 and wind tunnel Reynolds number of six million. Computational modeling revealed the source of a measured pressure peak on the projectile surface as the termination of a region of high pressure, low speed flow. Computational solutions also indicate the importance of three-dimensional, non-axisymmetric simulations over axisymmetric approximations, and viscous flow modeling over inviscid flow.

Future computations for the wind tunnel projectile/sabot model will include turbulence and separated flow modeling, grid refinement in the axial direction between the projectile and the sabot petals, and additional configurations for which the sabot petal is at angle of attack, including test cases for one sabot petal with flow splitter plates. Numerical simulations will be attempted using larger Reynolds numbers approaching free flight values, with corresponding refinements in the computational grid. Computational methods and technology proven using wind tunnel configurations can be applied to fielded projectiles such as 25mm long-rod penetrators and the XM900 projectile and sabot. Detailed 3D numerical computations can also be used to enhance rapid engineering sabot design codes that are based on local shock/expansion gasdynamic analysis techniques.¹⁵ Further, a subprogram to integrate surface pressure and shear stress is in development and will be used to determine the aerodynamic forces and moments acting on the sabot surfaces. This output could be used in quasi-steady 6 degree-of-freedom sabot trajectory simulations. The ultimate goal is a time-accurate, linked aerodynamic solution/trajectory simulation.

V. ACKNOWLEDGEMENT

The author wishes to acknowledge Dr. E.M. Schmidt, Chief of the Fluid Physics Branch, Launch and Flight Division, USA BRL, for his substantial fiscal support of this work.

VI. REFERENCES

1. Schmidt, E.M., and Shear, D.D., "Aerodynamic Interference During Sabot Discard," AIAA Journal of Spacecraft and Rockets, Vol. 15, No. 3, May-June 1978, pp. 162-167.
2. Schmidt, E.M., and Shear, D.D., "Launch Dynamics of a Single Flechette Round," US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, ARBRL-01810, Aug. 1975.
3. Schmidt, E.M., "Wind-Tunnel Measurements of Sabot-Discard Aerodynamics," AIAA Journal of Spacecraft and Rockets, Vol. 18, No. 3, May-June 1981, pp. 235-240.
4. Schmidt, E.M., and Plostins, P., "Aerodynamics of Asymmetric Sabot Discard," US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, ARBRL-MR-03281, June 1983.

5. Chakravarthy S.R., Szema K.Y., Goldberg U.C., Gorski J.J., and Osher S., "Application of a New Class of High Accuracy TVD Schemes to the Navier-Stokes Equations," AIAA-85-0165, Proceedings of the 23rd AIAA Aerospace Sciences Meeting, Reno NV., January 14-17, 1985.
6. Ota, D.K., Chakravarthy S.R., and Darling, J.C., "An Equilibrium Air Navier-Stokes Code for Hypersonic Flows," AIAA-88-0419, Proceedings of the 26th AIAA Aerospace Sciences Meeting, Reno NV., January 11-14, 1988.
7. Palaniswamy, S., and Chakravarthy, S.R., "Finite Rate Chemistry for USA Series Codes: Formulation and Applications," AIAA-89-0200, Proceedings of the 27th AIAA Aerospace Sciences Meeting, Reno, NV, January 9-12, 1989.
8. Baldwin, B.S. and Lomax, H., "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA-78-257, Proceedings of the 16th AIAA Aerospace Sciences Meeting, Huntsville, AL, January 16-18, 1978.
9. Goldberg U.C., "Separated Flow Treatment with a New Turbulence Model," AIAA Journal, Vol. 24, No. 10, October 1986, pp. 1711-1713.
10. Chakravarthy, S.R. "A New Computational Capability for Ramjet Projectiles," ARBRL-CR-595, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD, March 1988. (Also Nusca, M.J., Chakravarthy, S.R., and Goldberg, U.C., US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, ARBRL-TR-02958, Dec. 1988.)
11. Szema, K.Y., Chakravarthy S.R., Riba, W.T., Byerly, J., and Dresser, H.S., "Multi-Zone Euler Marching Technique for Flow Over Single and Multi-Body Configurations," AIAA-87-0592, Proceedings of the 25th AIAA Aerospace Sciences Meeting, Reno NV., January 12-15, 1987.
12. Szema, K.Y., Chakravarthy S.R., Pan, D., Bihari, B.L., Riba, W.T., Akdag, V.M., and Dresser, H.S., "The Application of a Unified Marching Technique for Flow Over Complex 3-Dimensional Configurations Across the Mach Number Range," AIAA-88-0276, Proceedings of the 26th AIAA Aerospace Sciences Meeting, Reno NV., January 11-14, 1988.
13. Woan, C.J., and Chakravarthy S.R., "Transonic Euler Calculations of a Wing-Body Configuration Using a High-Accuracy TVD Scheme," AIAA-88-2547, Proceedings of the 6th AIAA Applied Aerodynamics Conference, Williamsburg, VA, June 6-8, 1988.
14. Chakravarthy S.R., Szema K.Y., and Haney, J.W., "Unified Nose-to-Tail Computational Method for Hypersonic Vehicle Applications," AIAA-88-2564, Proceedings of the 6th AIAA Applied Aerodynamics Conference, Williamsburg, VA, June 6-8, 1988.

15. Siegelman, D., Wang, J., and Crimi, P., (AVCO Systems Division), "Computation of Sabot Discard," US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, Contract Report, ARBRL-CR-00505, February 1983.

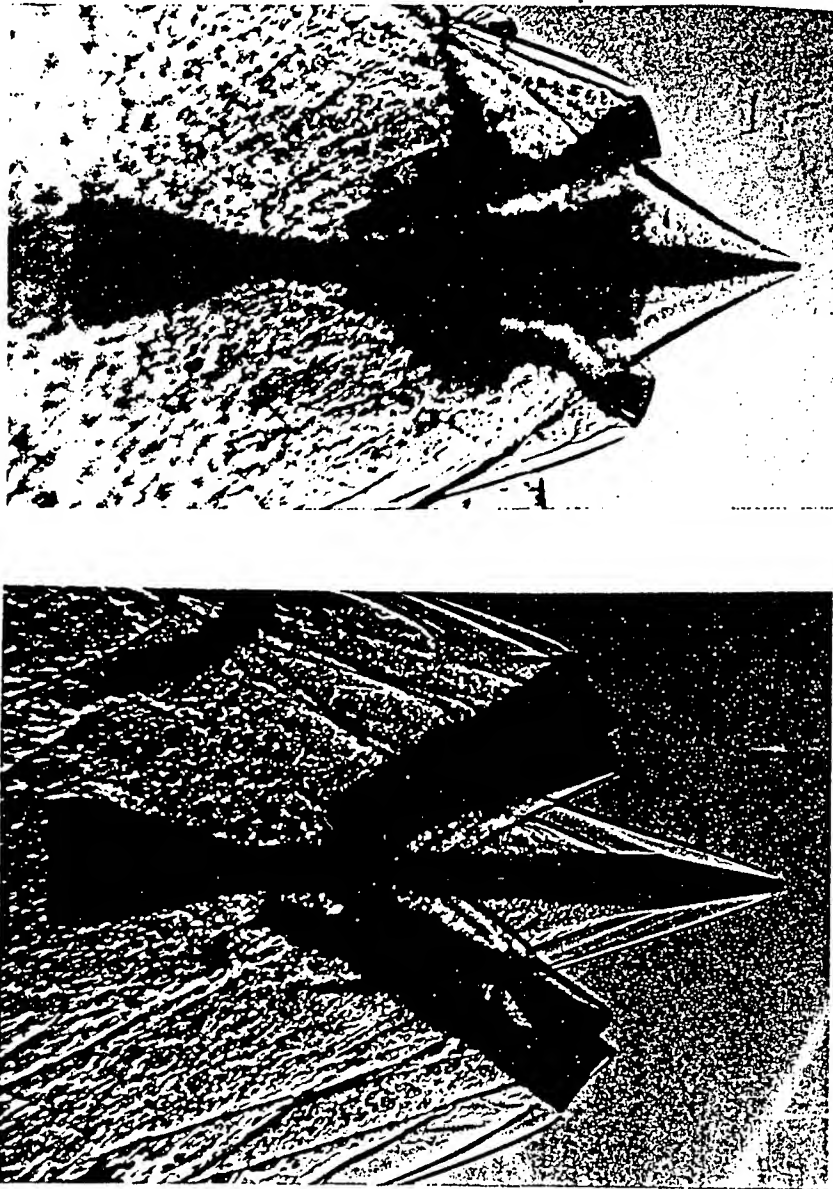


Figure 1: Spark Shadowgraphs of Typical Projectile/Sabot Discard Sequence.

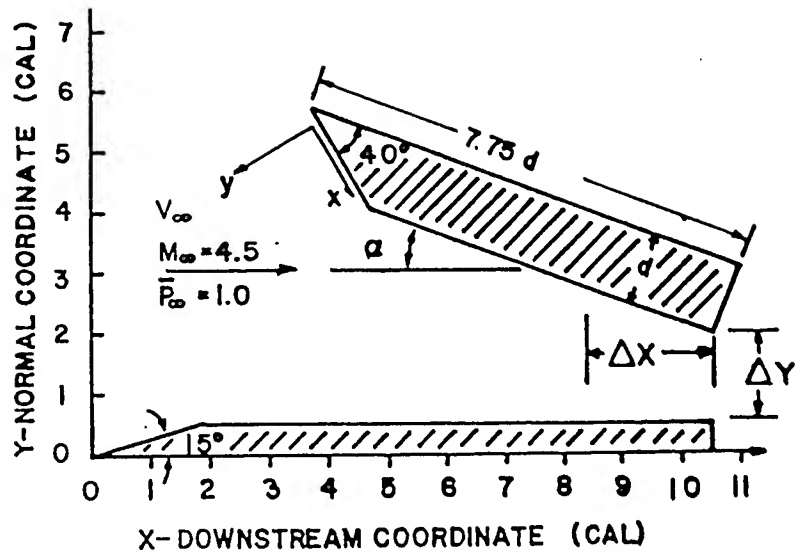


Figure 2: Schematic of Wind Tunnel Model and Nomenclature.

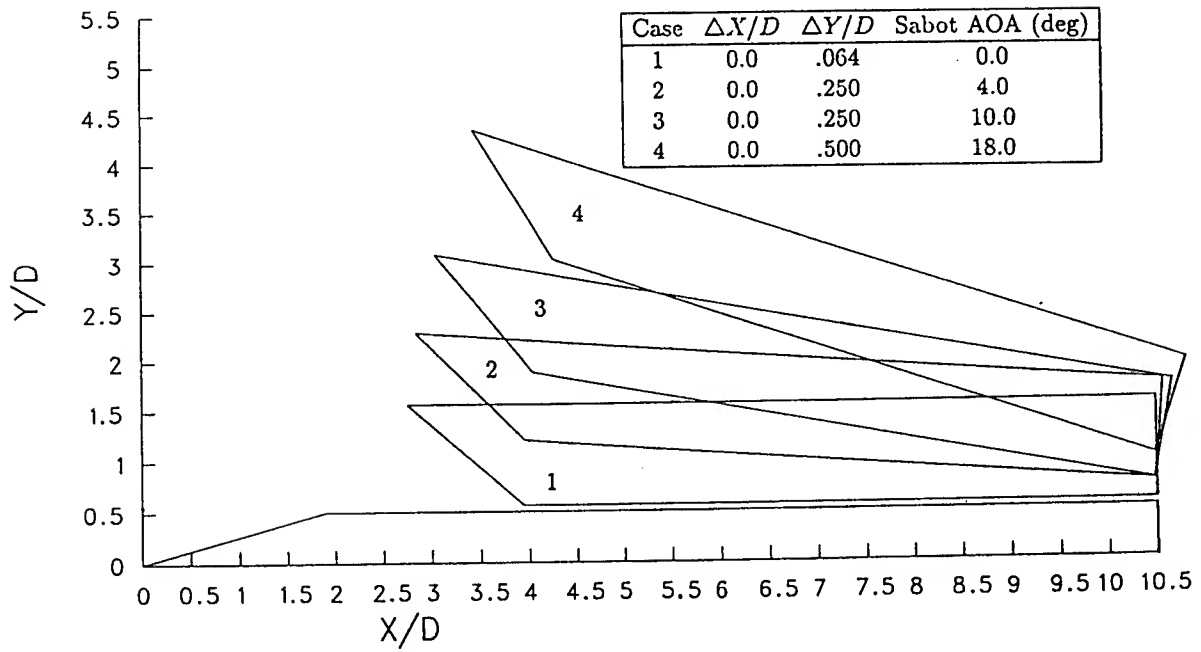


Figure 3: Schematic of Simulated Sabot Discard Wind Tunnel Test Sequence.

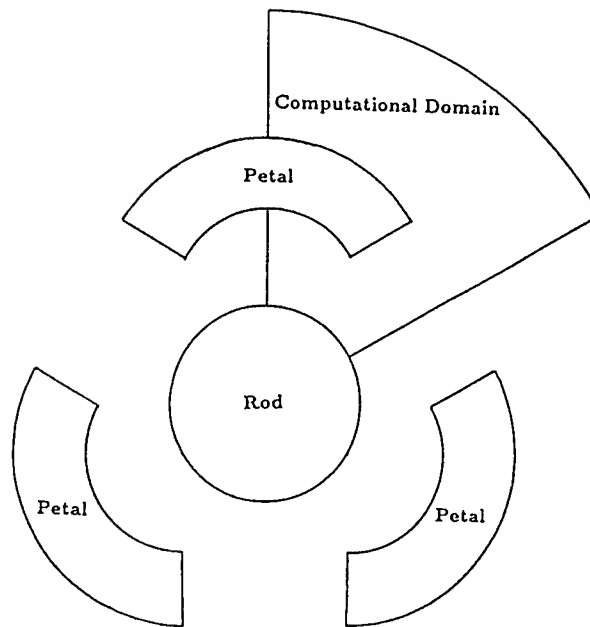


Figure 4: Schematic of Symmetric Sabot Discard. Computational Domain Shown.

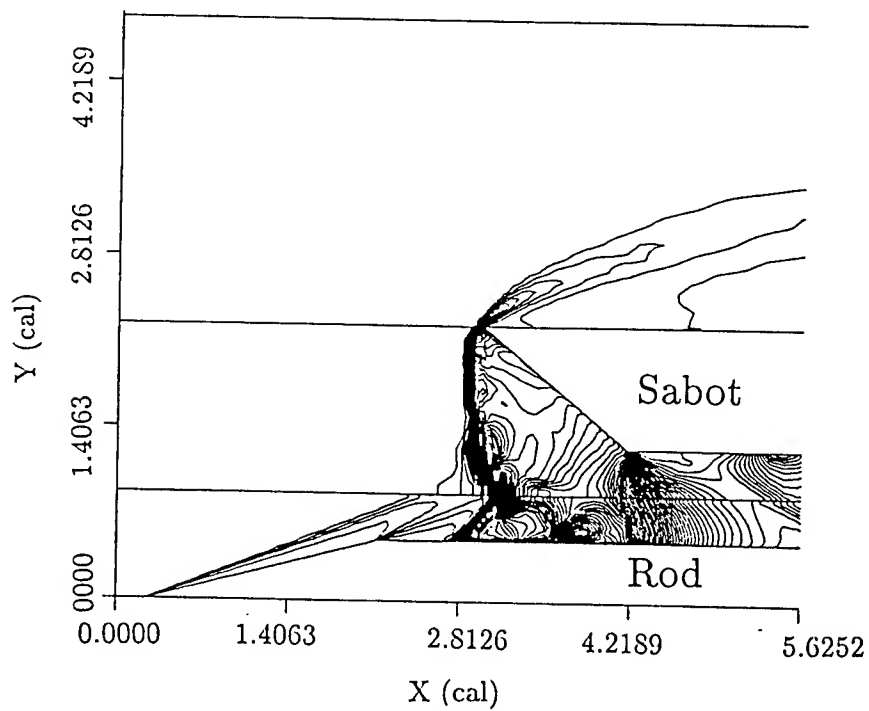


Figure 5: Pressure Contours, Laminar Flow, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

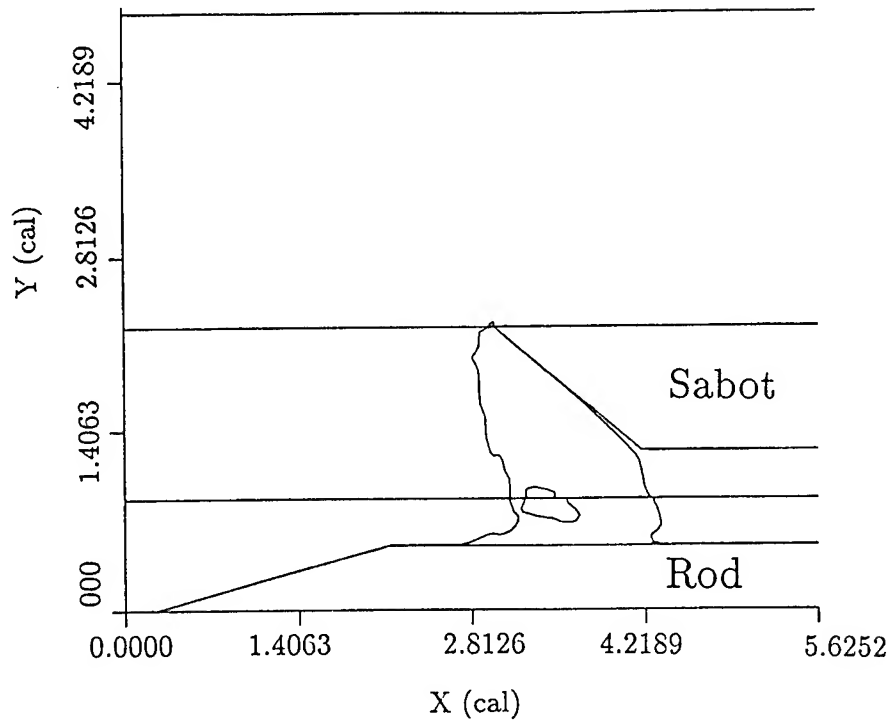


Figure 6: Sonic (Mach = 1) Contours, Laminar Flow, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

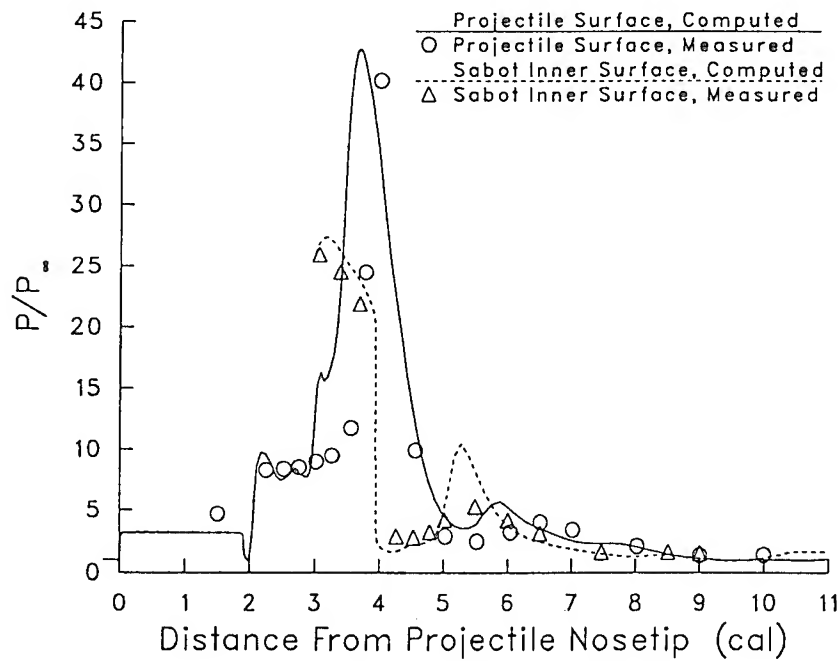


Figure 7: Pressure Distributions on the Projectile and Sabot Surfaces in the Pitch Plane ($\phi = 0$), Laminar Flow, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

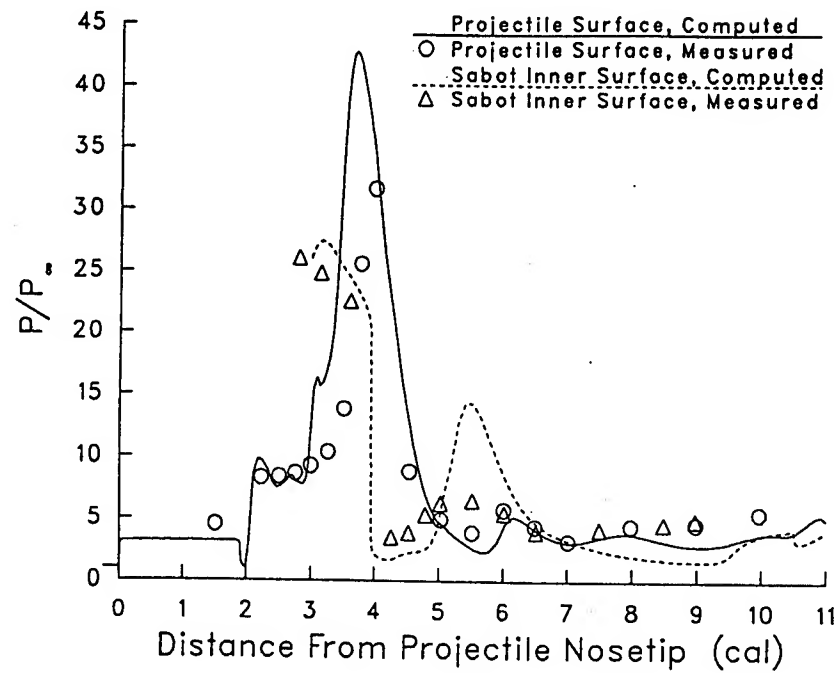


Figure 8: Pressure Distributions on the Projectile and Sabot Surfaces in the Pitch Plane ($\phi = 0$), Laminar Flow, $\Delta x/D = 0$, $\Delta y/D = .5$, Sabot at 4° Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

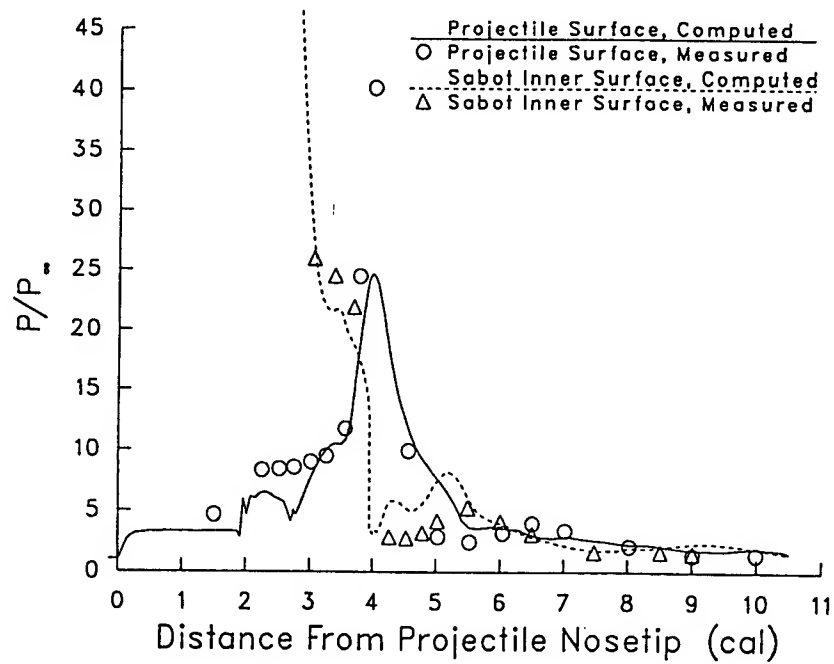


Figure 9: Pressure Distributions on the Projectile and Sabot Surfaces in the Pitch Plane ($\phi = 0$), Inviscid Flow, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

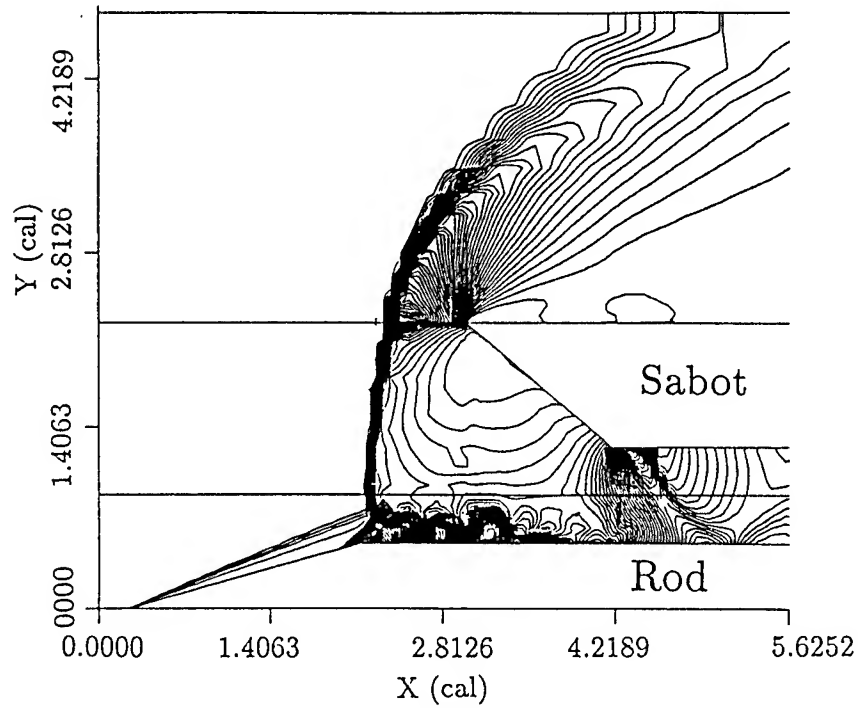


Figure 10: Pressure Contours for Axisymmetric Geometry, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

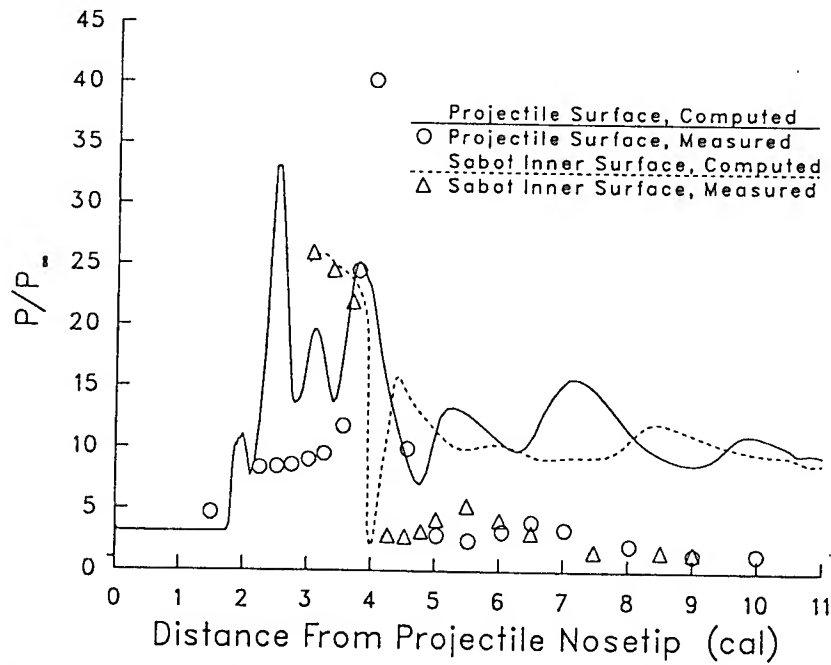


Figure 11: Pressure Distributions on the Projectile and Sabot Surfaces in the Pitch Plane ($\phi = 0$), Axisymmetric Geometry, $\Delta x/D = 0$, $\Delta y/D = .75$, Sabot at Zero Yaw, $M_\infty = 4.5$, $Re = 6.6$ million.

Predicting Target Detection Skill (U)

*LTC Timothy R. O'Neill

Ms Dee Ann Batten

Ms Susan Woontner

Department of Behavioral Sciences and Leadership
United States Military Academy
West Point, New York 10996

As the range of antiarmor weapons increases, the problem of visual target acquisition becomes more daunting. Most soldiers cannot detect targets at ranges beyond two kilometers, and visual recognition much beyond battlesight range is often chancy at best. Even with the fielding of thermal sights, this problem restricts the usefulness of weapons and calls into question the adoption of systems with effective ranges farther than the limited capacity of the soldier to service.

There are two approaches to this challenge in human performance: training and selection. Training is an attractive alternative because it does not cause pressures for the assignment of soldiers with specific talents to specific jobs; however, the evidence from human performance studies suggests that, though target detection skill can be improved by training, it is likely to plateau when some individual innate level is reached. In addition, soldiers with high natural talent are likely to benefit most from training.

Selection of soldiers with high potential for assignment to such specialties must be based on procedures that are reliable, valid, and efficient in application, and the predictive power of these tests must be strong enough to justify intervention in the assignment process.

The work described here was conducted at the Human Sciences Laboratory of the Department of Behavioral Sciences and Leadership, United States Military Academy, in 1988, under a grant from the Army Research Institute for the Behavioral and Social Sciences. The objective was the development of a simple, powerful and reliable test battery for identification of observers with high performance potential in tasks requiring complex visual search. We employed tests of two general visual faculties: visual lobe size and contrast sensitivity.

The Visual Lobe

Although a term common in engineering psychology, "visual

lobe" is somewhat misleading, since the implication is of an anatomical structure. Renshaw¹ described the concept as early as 1945, using the term "visual form field"; Christenson² and Smith³ used the term "working field of view." In essence, the visual lobe is an area surrounding the center of vision within which shapes are identifiable.

The visual lobe is associated with one of the two visual pathways: the geniculostriate, as opposed to the relatively primitive tectopulvinar, which serves peripheral as opposed to central vision. Schneider has described this collection of central vision pathways as the "what is it?" system, while the peripheral pathways deal with the question "where is it?"⁴

But the visual lobe cannot be based purely on the anatomical structure of the eye. For one thing, central (foveal) vision is much narrower than the empirically-derived visual lobe dimensions; and the visual lobe is to some extent subject to enlargement with training ("speed reading" exercises are built around this observation), and the retina's structure is not. Hence, we will regard the visual lobe as a construct rather than a physical entity.

The visual lobe has been shown in a variety of studies^{5,6} to be correlated with performance in visual search tasks. Most of these studies, however, have employed rather sterile laboratory paradigms, and their applicability to the battlefield case is unclear.

Contrast Sensitivity

It stands to reason that measures of visual acuity should be strongly correlated with visual search performance. Surprisingly, traditional measures - Snellen and Landolt Rings tests - have yielded disappointingly weak results.⁷ Ginsburg, for example, reports that Snellen acuity was not significantly related to target detection in a sample of Air Force pilots.⁸

Snellen acuity, however, measures only sensitivity to very high spatial frequencies. Years of research in pattern perception have demonstrated the existence of multiple discrete channels in the visual system selectively sensitive to rather narrow ranges of spatial frequency (size).⁹ These channels might, when directed at a patch of woods, consist of "forest detectors," "tree detectors," and "leaf detectors." Sensitivity (measured as the inverse of the contrast at threshold) varies across a range of frequencies, resulting in a spatial frequency modulation transfer function (MTF). Highest human sensitivity is in the range of 4-6 cycles per degrees of retinal angle, with weaker performance (less sensitivity) for very low (for example, .5 c/d) or very high (20-30 c/d). Observers' spatial MTF's vary considerably. We should

also note that the traditional Snellen-Landolt type test measures only the highest frequencies.

Ginsburg noted, in the same study that saw the failure of Snellen as a target detection predictor, that the spatial MTF was strongly correlated with target detection, with a coefficient in the .9 range.¹⁰ This and other diagnostic virtues of contrast sensitivity make it an attractive tool for vision research and a prime candidate for observer selection.

Method

Observers: Thirty volunteers from the Staff and Faculty, United States Military Academy, and the 1st Battalion, 1st Infantry, participated in this study. All observers had natural or corrected 20/20 vision (Snellen) and normal color vision (Dvorine pseudoisochromatic). Ages ranged from 19 to 42. All had some experience in target detection, through initial training or incident to MOS requirements.

Tests and Apparatus:

Contrast sensitivity was measured at 1.5, 3, 6, 12 and 18 cycles/degree using the Vistech VCTS test. This is a wall chart with systematic rows and columns of patches with sinusoidal gratings of various orientations and contrasts. Observers were required to identify the orientation of the grating for each patch; sensitivity was derived from the lowest contrast patch with an identifiable grating. As a backup, the observers were tested against a series of drifting gratings generated by computer on a phosphor screen.

The visual lobe measure yielded to a locally-designed test (Visual Lobe Test). This consisted of observation of a 19-inch video display viewed at a distance of 1 meter. A central fixation point appeared, and subsequently letters flashed for a duration of 250 ms each at varying distances from the center along the vertical and horizontal meridians. There are ten trials at each position, 240 total trials. The width of the visual lobe is derived from the distances from center at which an observer can correctly identify a letter with fifty percent accuracy (Figure 1)

In addition, we administered a Visual Search Test, in which observers viewed the display filled with a random matrix of identical letters, with one different target letter (a "T" among "O"'s, for example). The system measured elapsed time from presentation to detection, signalled by closure of a finger switch.

Finally, observers completed the Degraded Letters Test, a

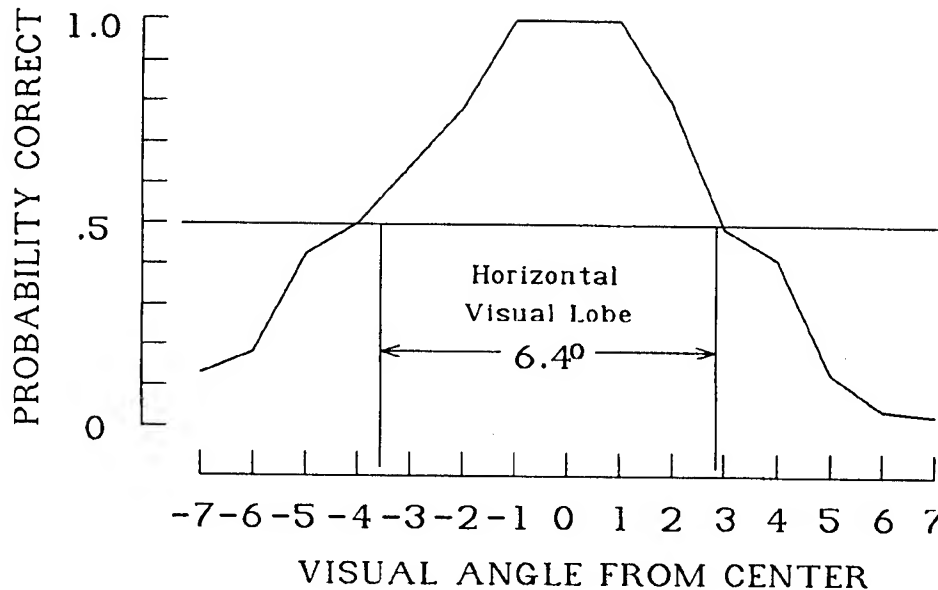


Figure 1: Typical visual lobe measure

paper measure correlated with contrast sensitivity developed at West Point in the 1970's.¹¹ This consists of a series of dot-matrix letters with random additions and deletions, in effect "camouflaging" the shapes. (Figure 2)



Figure 2: DLT test figures.

Target detection skill was measured using three methods.

First, observers viewed a series of 35mm transparencies rear-projected on a ground glass screen. Each slide showed a field site with a hidden vehicle, target size less than 3 degrees visual subtense. The system measured response time from stimulus presentation to switch closure signalling detection. Target detection was verified using a superimposed grid.

Second, the first ten slide presentations were accompanied by measurement of eye movement patterns using an ASL 1996 eye view monitor. After analysis, it was possible to determine the number of fixations necessary to find the target, and the mean number of fixations for ten targets was added to the mean visual response time.

Finally, observers viewed a series of slides containing a small silhouette of a tank superimposed on a field of visual noise (Figure 3). Stimuli were presented for 250 msec using a projection tachistoscope. The observer reported in each case (1) whether a target was present, and (2) its position on a nine-section grid. Twenty-five percent of the stimuli were "null trials" with no target. The result was a signal detection paradigm in which probability of detection and probability of false alarm are critical.

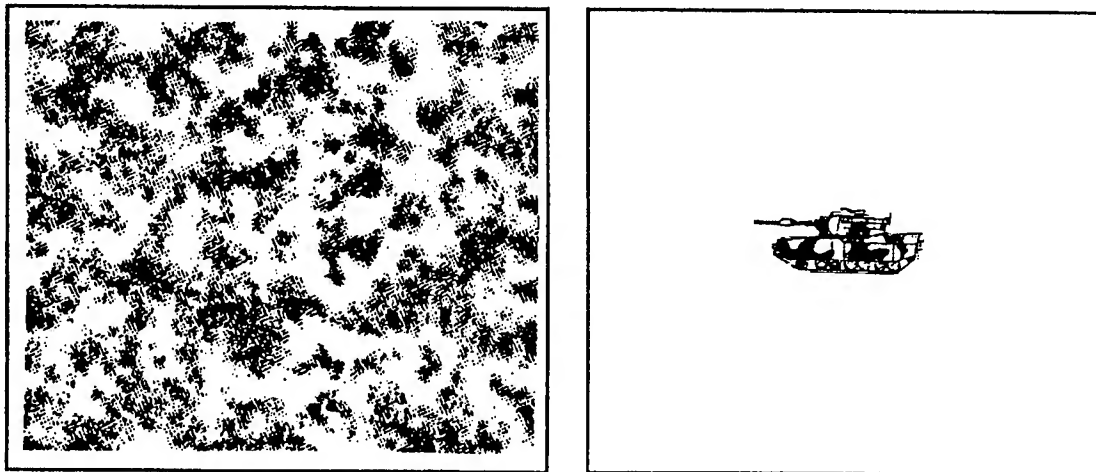


Figure 3: Noise pattern (left); target shape (right). The target shape has been significantly enlarged.

Results

Table 1 shows the correlation matrix. C1 through C5 represent contrast sensitivity at 1.5, 3, 6, 12 and 18 c/d respectively; DL is the Degraded Letters Test; VL is the Visual Lobe Test; PH is probability of detection in the signal detection test; MD is mean detection time; and MF is mean number of fixations per trial.

Note that contrast sensitivity at 6, 12, and 18 c/d were strongly correlated with visual lobe size (Pearson $r = .5$, $.69$, $.55$ respectively) This is consistent with Ginsburg's results. These measures were also highly correlated with probability of detection in the signal detection task ($.58$, $.48$, $.35$). Mean detection time was predicted strongly by a number of variables, but most notably higher contrast sensitivity (12 c/d). However, DLT and VST were also moderately correlated (both with Pearson r of $.54$) with detection time,

TABLE 1: CORRELATION MATRIX

	C1	C2	C3	C4	C5	DL	VS	VL	PH	MD	MF
C1	1	.31	.43	.23	-.03	.06	-.11	.23	.21	.19	.16
C2		1	.67	.38	.18	.07	-.15	.07	-.09	.06	-.20
C3			1	.67	.57	.02	-.42	.50	.58	-.57	.00
C4				1	.60	-.19	-.43	.69	.48	-.67	-.38
C5					1	-.09	-.48	.55	.35	-.46	.16
DL						1	-.09	-.14	.08	.54	.08
VS							1	-.35	-.19	.54	.08
VL								1	.66	-.68	-.15
PH									1	-.65	-.39
MD										1	.11

and virtually independent of each other. Mean fixation time was moderately correlated with contrast sensitivity at 12 c/d and, as we might expect, with probability of detection.

We felt that the three measures of target detection should be viewed together, since they represented subtly different aspects of the actual task; hence we used a canonical correlation of all predictor variables on all dependent measures.

The result was a canonical correlation coefficient of .897 — an exceptionally strong predictive power, in essence explaining roughly eighty percent of the variance in detection performance.

Discussion

The principal objective of devising a test battery for observer selection was clearly met. The four tests include one recently available and increasingly widespread in the practice of optometry and oculometry; two that require only a personal computer; and one written test requiring about five minutes to complete.

But there were surprises in the results that may prove more valuable than the test battery. The most important was the discovery of a small population of "super-lobe" observers (Figures 4 to 6). These soldiers had, essentially, no limitation to apparent foveal vision. This is anatomically impossible, and must be related to some other phenomenon. We note that the "super-lobes" also had exceptionally high spatial MTF's. One possibility is that their contrast sensitivity is so excruciating that they were able to fixate on the fading phosphor trace of the target letter! This hypothesis will be

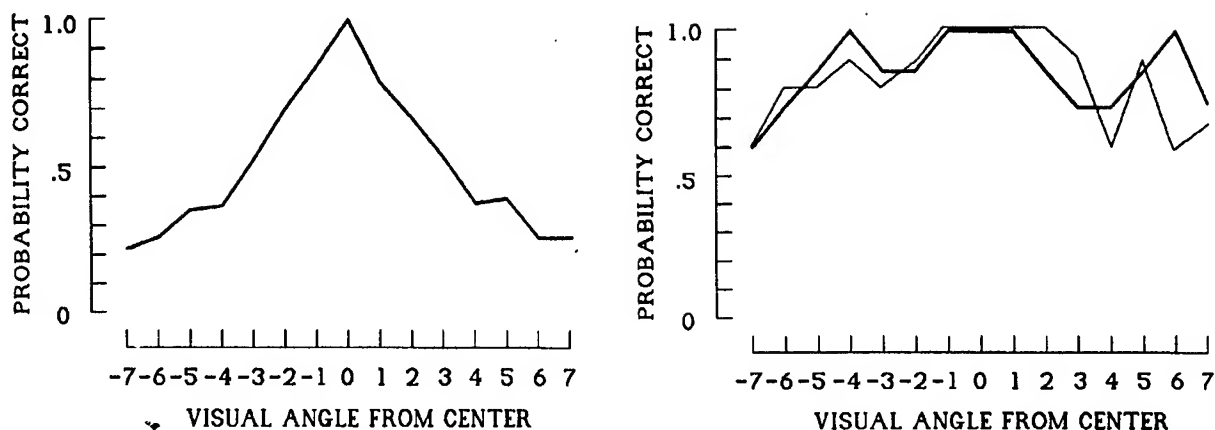


Figure 4: Atypical visual lobes; aggregate visual lobe distribution (left), two "super-lobe" examples (right)

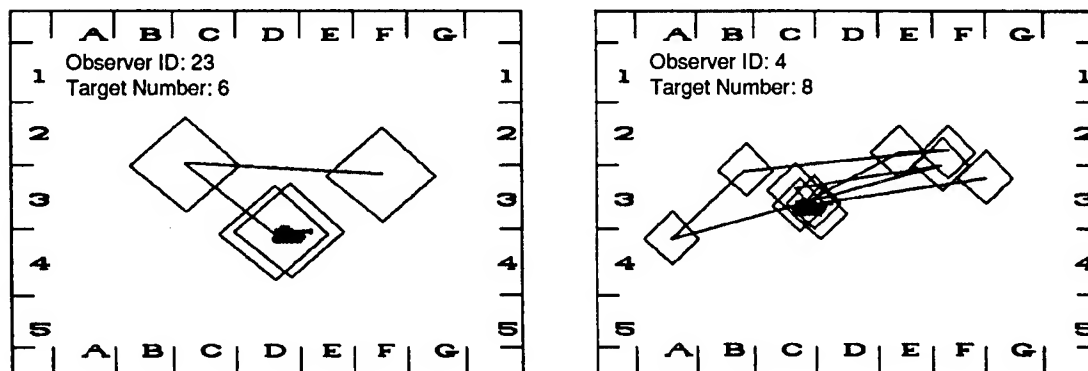


Figure 5: Typical search patterns. The diamond shapes are based on the measured horizontal and vertical visual lobes for the observers.

tested using a tachistoscopic method.

Whatever the details of the visual faculties that contributed to observer performance, the results of this stage of the program show that prediction of target detection skill is now simple, accurate and powerful, and yields collateral insights into the incredible complexity of vision.

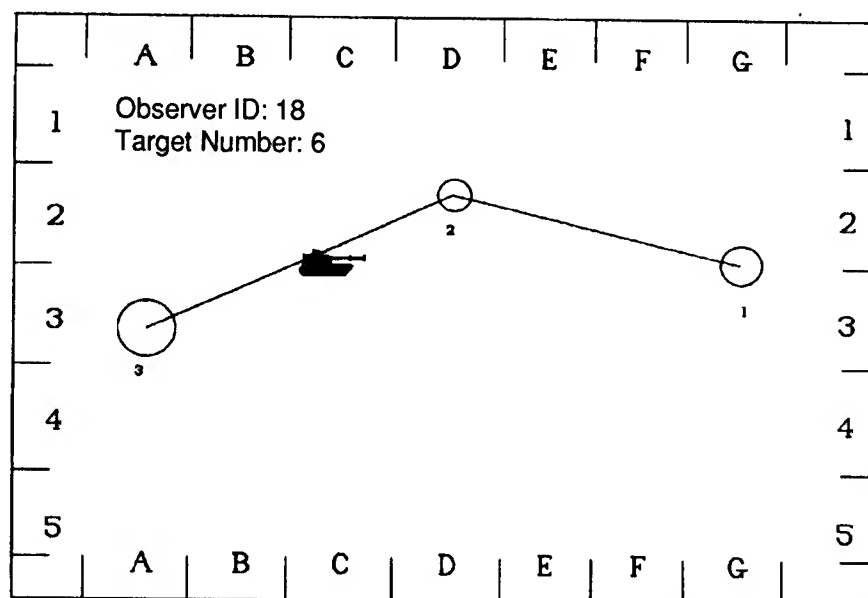


Figure 6: "Super-lobe" scan. In this case, circles represent fixations, radii are proportional to dwell times. Note that none of the fixations falls close enough to the target to explain detection.

Endnotes

1. Renshaw, S. (1945) The visual perception and reproduction of forms by tachistoscopic methods. Journal of Psychology, 20, 217-232.
2. Christenson, J. and Crannell, C. (1955) The effect of selected visual training procedures on the visual form field. WADC-TR-54-239, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.
3. Smith, S. (1961) Time required for target detection in complex abstract visual displays. Memorandum 2900-235-R, Institute of Science and Technology, the University of Michigan.
4. Schneider, G. E. (1969) Two visual systems. Science, 163, 895-902.

5. Bellamy, L. and Courtney, A. (1981) Development of a search task for the measurement of visual acuity. Ergonomics, 24, 497-509.
6. Gallwey, T. (1982) Selection tasks for visual inspection on a multiple fault type task. Ergonomics, 37, 1077-1092.
7. Wiener, E. (1975) Individual and group differences in inspection. In Drury, C. and Fox, J. (Eds.) Human Reliability in Quality Control. London: Taylor and Francis, 101-102.
8. Ginsburg, A. and Evans, D. (1982) Contrast sensitivity predicts pilots' performance in aircraft simulators. American Journal of Optometry and Physiological Optics, 59, 105-109.
9. Blakemore, C. and Campbell, F. (1969) On the existence in the human visual system of neurones selectively sensitive to orientation and size of retinal images. Journal of Physiology (London), 23, 153-164.
10. Ginsburg, A. and Evans, D., op. cit.
11. O'Neill, T., Johnsmeyer, W., Brusitus, J. and Taylor, D. (1977) Investigation of psychometric correlates of camouflaged target detection and recognition. West Point, NY: Office of Military Leadership.

OPTICAL STUDIES OF THE FLOW START-UP IN CONVERGENT/DIVERGENT NOZZLES

*Klaus O. Opalka

Dixie M. Hisley

Terminal Ballistics Division

U.S. Army Ballistic Research Laboratory

Aberdeen Proving Ground, MD, 21005-5066

A shock tube test was carried out to investigate the influence of the divergent cone angle and length of four convergent/divergent nozzles on the pressure signature in the expansion region downstream from the nozzle. The diaphragm was mounted in the throat of these nozzles and located at the upstream edge of the test section window of high-quality optical glass through which shadowgraphs and schlieren pictures could be taken. The objectives of this investigation were to obtain optical records of the shock formation and flow start-up processes in the divergent nozzles immediately after rupturing the diaphragm, and to facilitate comparisons with one- and two-dimensional hydrocode computations. The optical measurements were complemented by measurements of the static pressure at three locations downstream from the nozzles. The test set-up and the results of this experimental study are presented and the most significant findings discussed.

1. BACKGROUND

The U.S. Army has been developing the concept of a Large Blast and Thermal Simulator (LB/TS) suitable to test full-scale military equipment for its vulnerability to thermal and blast effects of decaying blast waves such as are associated with nuclear explosions. Large blast simulators are basically shock tunnels whose cross-sectional areas vary along their lengths. They employ convergent/divergent nozzles at the driver exits to retard the outflow of the high-pressure driver gas and generate long flow durations because straight shock tubes cannot produce the flow durations which are observed in decaying blast waves. The diaphragms are mounted in the throats of the nozzles. A single driver is technically impractical because of the size of these facilities and the required supply pressures in the drivers. A number of smaller drivers have to be employed to accommodate the necessary initial test conditions.

The U.S. LB/TS concept is described in detail in the References^{1-3*} and illustrated in Figure 1. Blast waves will be simulated by releasing compressed gas from several high-pressure steel driver tubes into a large expansion tunnel constructed of prestressed concrete.

*References are listed at the end of the report.

The test section for the targets is located in the expansion tunnel, seven diameters downstream from the exit of the driver nozzles. It has a semi-circular cross-section with a 10m radius. This size is deemed necessary to avoid blocking of the flow about the target. The thermal simulation will be effected through aluminum/oxygen combustion near the target.

Each driver of the planned U.S. LB/TS facility has the diaphragm which separates the high-pressure from the low-pressure region mounted in the throat of a convergent/divergent nozzle. When the diaphragm is ruptured, a shock forms and travels downstream into the divergent section, while a rarefaction fan travels upstream into the convergent section. The divergent nozzles increase the thrust on the reaction pier which is generated by the "rocket motor" formed by the driver-nozzle combination. Therefore, eliminating the divergent nozzles would have the double advantage on the design and performance of the U.S. LB/TS of reducing the thrust on the pier as the drivers empty, as well as reducing the construction cost of the drivers. On the other hand, results of small-scale pilot experiments seem to indicate that there are pressure losses connected with the introduction of a large area discontinuity at the exit of the nozzle throat, which are lessened when divergent nozzles are used. Therefore, the use of divergent nozzles cannot be easily dismissed. However, if divergent nozzles are included in the U.S. LB/TS design, the nozzle length becomes a critical material parameter which itself depends on the cone half-angle chosen for the nozzle design.

Convergent/divergent nozzles have been extensively investigated, both in theory and experiment, and a summary of these efforts is presented, e. g., by Amann.⁴ Amann has shown in his research that the start-up process of the nozzle is shortest when the cone half-angle of the divergent nozzle is largest. The start-up process of the nozzle is defined as the time from the rupturing of the diaphragm to the establishment of a quasi-steady flow pattern in the nozzle. Consequently, a 90° angle at the exit of the throat should be optimal. However, the scope of Amann's research effort was limited to cone half-angles of 5° - 15° and it is not known whether the trend continues at the higher angles. Furthermore, in all of the known investigations, the diaphragm separating the high-pressure from the low-pressure region was located far upstream from the nozzle under investigation so that a well-formed shock would enter the nozzle from the upstream side. This is not the case in the present nozzle design, which has the diaphragm mounted in the throat of the nozzle. No previous experimental research is known for either the starting process, or the quasi-steady flow phase of this problem and, therefore, the present experimental study was initiated.

2. EXPERIMENTAL APPARATUS

The proposed experiments were carried out under contract at the Ernst-Mach-Institut (EMI), Freiburg im Breisgau, West Germany. The experimental apparatus at the EMI consists of a shock tube (which was modified for use in this study), a 24-spark Craz-Schardin camera and shadowgraph arrangement, and electronic control and recording equipment.

The EMI shock tube has a circular cross-section with an inner diameter of 200 mm in its most used configuration. The driver section is 1800 mm long. The driven section measures 8880 mm (Figure 2). Cellulose Acetate sheets (Ultraphan) have proven to be a good diaphragm material because they become brittle when stretched and have a high breaking velocity. The shock tube has a rectangular test chamber with a cross-section of 40 mm x 110 mm

located 7210 mm downstream from the driver section. The test chamber is followed by a 1570-mm long extension to the closed end. The side walls of the test chamber are formed by plane parallel windows (200 mm x 110 mm) of high quality optical glass to facilitate optical investigations of the shock wave propagation. An inlet channel of the same shape and cross-section as the test chamber, formed of polished aluminum plates, protrudes upstream into the axisymmetric driven section. A rectangular section is cut out of the arriving shock wave of circular shape by the sharp inlet edges of the channel walls (cookie cutter principle).

To carry out the present study, the existing shock tube had to be modified. The modifications included a new driver pressure chamber which could be mounted inside the inlet channel of the existing shock tube, and the various model parts for the divergent nozzles. The new driver pressure chamber included the convergent section and that part of the throat of the nozzle which lies upstream of the diaphragm (Figure 3a). The inner cross-section of the driver was 64 mm x 30 mm converging to a cross-section of 16 mm x 30 mm at the diaphragm. The length of the driver pressure chamber was chosen to 1000 mm in order to delay the arrival of reflections from the end wall. The diaphragms were composed of two Ultraphan sheets, each 0.1 mm thick, with an ignition wire of 0.08-mm diameter sandwiched in between them. With this construction, the diaphragm could be burst at an exactly predetermined moment in time by exploding the wire with an electric impulse.

Four models of convergent/divergent nozzles were fabricated, differing in the lay-out of the divergent section as follows:

- a) A convergent nozzle and throat section with the area ratios of the U.S. LBTS, but without the divergent section, also referred to as the 90° divergent nozzle (Figure 3b).
- b) An extended convergent/divergent nozzle with a divergent 45° cone half-angle, extending to the full height of the expansion tunnel (Figure 3c).
- c) A convergent/divergent nozzle with the area and length ratios and the cone half-angle (6°) of the nozzles used in the CEG-LBS (Figure 3d).
- d) An extended convergent/divergent nozzle with the area and length ratios of the Q1D-computational LB/TS model, in which the nine driver nozzles of the U.S. LB/TS were combined into one nozzle. This divergent section has a cone half-angle of 16° and its length is extended to the full height of the expansion tunnel (not shown).

The test objectives required locating and mounting the diaphragm in, or near, the test section window of the EMI shock tube in such a way that the view of the downstream region of the throat and divergent nozzle was only minimally obstructed. The lengths of the various divergent nozzle configurations make it necessary to relocate the models in the test section window so that optical records may be taken of all waves of interest from one to three diameters downstream from the diaphragm. This requirement made it necessary to locate the diaphragm either at, or at a multiple of one to three times one half of the window length (100 mm) upstream from the left edge of the test section window.

The four model configurations were tested at seven driver pressure ratios ranging from 4 to 188. These driver pressure ratios are related to certain simulated blast wave conditions in the test section of the U.S. LB/TS expansion tunnel, which were defined with reference to shock strength, duration, and static and dynamic impulse. For the purpose of this study, the initial shock strength in the throat of the nozzle was determined by iteratively solving the basic

shock tube equation. To test the functionality of the modified shock tube and to get a feel for the starting process of the various nozzles, the models were tested at the lowest specified driver pressure level ($P_{41}=4$) first. Tests then continued with the next highest driver pressure level ($P_{41}=14$). The specified operating pressure was 1300 kPa (13 bar) over atmospheric pressure, except for the lowest driver pressure ratio where the driver pressure was 300 kPa (3 bars) over atmospheric pressure. Then, the driven section was evacuated as necessary until the desired driver pressure ratio was established.

A 24-frame Cranz-Schardin spark camera and shadow/schlieren system was the principal tool for the optical study. The shadow method is well suited for the visualization of shock waves; the density gradients in the flow are made visible by the schlieren method. In the following experiments, both methods were combined and, in this way, clear images were obtained of the shock waves as well as of the density gradients. Other equipment used included the trigger gages and electronic delay units for the timing of the Cranz-Schardin camera and several 12-bit transient recorders. Kistler quartz pressure gages of type 603B were used to sense the static pressure. They have a characteristic frequency response of approximately 500 kHz. Electrical filters with a maximum frequency response of 180 kHz were installed between the pressure gage and the charge amplifier in order to prevent an overshoot of the signal.

The Toepler structure using sharp knife edges was found to yield the best results for the schlieren method. For the shadow method to work well, a point source of light is required. The 24 light sources in the Cranz-Schardin camera have a diameter of less than 1 mm. The sparks from these point sources are focussed onto 24 objective lenses of the camera by means of a concave mirror. The objective lenses are focussed on a reference plane typically at a distance of 40 cm from the center plane of the test chamber. At the moment a spark is generated, it is projected through the objective lens onto which it is focussed onto the film plane. The advantage of this procedure is that there are no moving parts aligning the path of the light rays. Therefore, the optical resolution is determined essentially by the aperture of the objective lenses. The effect of parallax is minimal in the present set-up where a concave mirror with a long focal length ($f = 350$ cm) is used in conjunction with a test section of shallow depth (4 cm).

The test was prepared by mounting the desired model configuration and the diaphragm. The shock tube was then closed, the Cranz-Schardin camera and the pressure recorders were readied, and all electrical connections were tested. The firing and recording controls were activated in the darkened room, and the recording film holder was opened before the driver was pressurized. The rupturing of the diaphragm was initiated with an electric pulse of 10-20 kV to explode the wire embedded between the two Ultraphan layers. The same pulse triggered the camera and the other recording equipment after some preset delay time. The preset delay time between the ignition pulse and the first spark pulse to the Cranz-Schardin camera was measured by a time counter. Some interference from the strong trigger pulse was suffered on the recording equipment, and a few pressure records were lost. A test log was kept to record the initial test and ambient conditions, including the trigger times and the spark frequency for the Cranz-Schardin camera and for the pressure recorders. The delay time between the bursting of the diaphragm and the first spark of the camera was additionally measured by a time counter.

3. RESULTS AND DISCUSSION

With each test shot, nineteen shadowgraphs and five schlieren pictures were recorded on film by the Cranz-Schardin camera at constant time intervals. The time intervals were estimated from the expected shock and particle velocities, based on an analytical solution of the shock tube problem applied to the throat channel of the nozzles for each of the specified driver-pressure ratios, and bracketed by the associated sound velocity expected in the expansion tunnel downstream from the nozzles.

The 16° divergent half-nozzle is suited best for an explanation of the flow start-up phenomena because the available viewing area (400 mm) is the largest among the configurations tested. A sample sequence of shadowgraphs for the 16° divergent half-nozzle at a driver-pressure ratio of 14 (Code N16/014) is shown in Figure 4. Two test shots were put together to present the whole half-nozzle. The heavy vertical line halfway through the nozzle indicates that the two photographs needed to record the flow throughout the entire half-nozzle at a particular point in time did not overlap. Flow phenomena like the incident shock, the contact surface, regions of flow separation, and a system of three recompression shocks can be seen developing and moving past the observation area. The low pressure case was chosen because all the flow phenomena of interest remain inside the viewing area until the start-up process of the nozzle is completed. The flow phenomena of interest are described in the following paragraphs.

The **incident shock** begins to form immediately after the flow is initiated by bursting the diaphragm. Since the diaphragm breaks into fragments of uneven sizes, the opening process is asymmetrical and causes an uneven flow progress across the height of the throat. Initially, the shock consists of a system of two to three curved shocks (106 μ s) that reflect from the ceiling or the floor of the throat channel and gradually combine into one shock (206 μ s). This formation period takes about 400 μ s, by which time the shock has progressed well into the divergent section of the half-nozzle (406 μ s). A shock reflection is generated at the ceiling of the half-nozzle exit (822 μ s) which proceeds across the flow area and is itself reflected later on at the floor plate of the expansion channel.

The **contact surface** between the expanding driver gas and the gas in the low-pressure region of the divergent half-nozzle follows the incident shock at close range. It can be easily detected because the associated flow turbulence it brings with it changes the background in the photographs. Because of the turbulence, the contact front is very rugged. Often, darker strings perpendicular to the flow direction can be detected in the turbulent region behind the contact front. It is surmised that these strings may represent compression waves which at later times are condensed into a recompression shock. Once the contact surface has entered the divergent section of the half-nozzle, a variety of phenomena begin to occur simultaneously.

Throat shocks appear inside the throat channel in an interplay between boundary-layer growth and flow acceleration (406 μ s and later). As many as five shocks have been observed temporarily. Two of these warrant special mention because of their general presence for almost all test conditions. The narrowest cross section exists at the diaphragm location. Once the flow reaches sonic condition in this cross section, it expands supersonically upon exiting into the throat channel and a standing X-type shock appears between 10-15 mm downstream from

the diaphragm location (606 μ s). A second stationary shock appears in the throat channel at about 50-60 mm downstream from the diaphragm location and remains there throughout the start-up process of the nozzle (822 μ s and later).

At the exit of the throat channel, an **expansion fan** begins to develop at the corner which the throat channel forms with the divergent section (822 μ s and later). Inside the divergent half-nozzle, a **recompression shock** begins to form (606 μ s). This shock develops into a bottle shock spawning a second shock ahead of itself (1.322 ms). The second shock develops into a bottle shock also, spawning a third shock ahead of itself in the process (1.822 ms). **Extensive flow separation regions**, at both ceiling and floor of the divergent half-nozzle, develop during this process. The start-up process of the nozzle comes to completion when this **system of bottle-type recompression shocks** becomes stationary either inside the nozzle, as in the present case, or downstream from the divergent nozzle for high driver-pressure ratios.

The shadowgraphs and schlieren pictures were evaluated by measuring the horizontal distances of the various flow phenomena from a reference point in the photograph, and recording them with reference to the diaphragm location ($X=0$) as functions of time beginning with the bursting of the diaphragm ($t=0$). From these records, x - t diagrams were generated. The test case described above was modelled in a quasi-one-dimensional context and simulated with the BRL-Q1D hydrocode.⁵ The flow phenomena observed in the nozzle start-up process were also found in the computational simulation. The progression with time of the incident shock, the contact surface, and the various recompression shocks were recorded, and the resultant data were plotted in x - t diagrams together with the results from the experimental study.

Figure 5 presents a comparison of experiment and computation in a x - t diagram for the case N16/014 discussed above. The distance from the diaphragm location (X) of the described flow phenomena is plotted as a function of the elapsed time from the burst of the diaphragm (t). The data points for the downstream nozzle position C (200-400 mm) were matched with the data points for the upstream nozzle position B (0-200 mm) by adding a time differential to the 24 recorded spark times so that the data points for the incident shock in position C would continue on a straight line drawn through the data points for the incident shock in position B. The computational data are represented as lines.

The diagram shows good agreement between the computational and experimental data for the incident shock and the contact surface. The data for the recompression shocks agree qualitatively. But instead of one ideal recompression shock, the real flow forms a system of several recompression shocks. From the diagram, one can learn that the velocities of the incident shock and of the contact surface become constant after a short acceleration period as the flow expands in the nozzle, then begin to decrease. This means that the flow in the nozzle is accelerating during the start-up process. After about 2.2 ms, the system of three bottle-type recompression shocks and the throat shocks have become stationary. The start-up process of the nozzle is now complete. At later times the process will reverse itself until the driver section has emptied and the pressure in the tube has equalized.

Experimentally, the technique of modelling one half of a symmetrical nozzle was necessary to obtain as large an expansion of the flow as possible in the existing facility. But in regard to the physical flow conditions, this technique was only partially successful. Figure 6

presents a comparison of the experiment with a two-dimensional hydrocode, Blast2D, computation. Two snapshots in time of the flow start-up process in the 45° half nozzle are shown, i.e., at 0.44 ms (Figure 6a) and at 1.44 ms (Figure 6b) after flow initiation. The shadowgraphs show that flow symmetry is preserved until the contact surface arrives. Behind the contact surface, boundary layers form along the bottom plate and the ceiling causing flow separation and destroying the symmetry of the flow field in the half-nozzle. Thus the comparison is disabled. The two computational density contour plots shown in Figure 6 are in reasonable qualitative agreement with the shadowgraphs. They show the same flow features as does the real flow, noticeably at the early time (Figure 6a).

Computationally, it has been common practice to model one half of a two-dimensional nozzle and plot the mirror image of the results across the symmetry line for a full-nozzle image. This technique has produced good qualitative agreement between full-nozzle experiments and inviscid computations in the past. In similar experiments by Amann,⁴ e.g., where a full-nozzle instead of a half-nozzle was used, the qualitative agreement between the results from a Blast2D computation⁶ and the experimental records were excellent. The differences between computation and experiment, particularly in the late-time comparison (Figure 6b), can be attributed to the fact that the lower wall is modeled as a symmetry boundary in the computation, while in the experiment the real flow separates from the lower wall due to the viscous nature of the real gas. For the same reason, the bottle shock is stretched further in the nonviscous computation.

For the 45° divergent half-nozzle, a viewing area of 300 mm was available which allowed us to observe the start-up flow phenomena for a longer period of time and distance behind the nozzle exit than for the other half-nozzles. The start-up process of the nozzle, which was defined in Section 1 as the time from the rupturing of the diaphragm to the establishment of a quasi-steady flow pattern in the nozzle, could be determined from the x-t diagrams for the 45° half-nozzle as a function of the driver-pressure ratio and is shown in Figure 7. For the 6° and 16° half-nozzle configurations, the flow start-up periods are expected to be somewhat larger, and for the 90° half nozzle configuration, very close to and below the values of the 45° half-nozzle.

The 90° half-nozzle is a nozzle without a divergent section. This configuration is especially interesting to the project engineer because the elimination of the divergent nozzles from the LB/TS design would save a great deal of material cost. The most significant flow feature of the 90° half-nozzle is that the flow forms its own divergent section once the local disturbances caused by the shock reflections from the ceiling and the floor have died down. The flow expansion angle was determined by reading and averaging the maximum and minimum angles of the flow expansion fan in the shadowgraphs and schlieren pictures, and was then plotted versus the driver-pressure ratio for the 90° half-nozzle in Figure 8. It was expected that the expansion angle would monotonously increase with the driver pressure; however, Figure 8 shows a maximum for the flow expansion angle at a driver-pressure ratio of $P_{41} = 110$ and a slightly lesser angle at the highest driver-pressure ratio, $P_{41} = 188$.

4. CONCLUSION

Four convergent/divergent nozzle configurations were successfully tested at seven driver-pressure ratios to study their flow start-up process. However, the viewing area for the 6° divergent half-nozzle and the 90° half-nozzle downstream from the nozzle exit was too limited, and additional optical test records are desirable for a comparative study of the flow start-up process downstream from the exits of the 90° and 6° half-nozzles. The present optical flow experiments have led to the following conclusions:

- a) The flow is very sensitive to local disturbances, e.g., flow asymmetries caused by the breaking of the diaphragm, diaphragm fragments and boundary-layer build-up along the bottom plate of the test chamber and the ceiling, and generates a system of recompression shocks in the throat channel and in the divergent section of the half-nozzle rather than a single recompression shock.
- b) For the half-nozzle experiments shown here, the flow along the lower wall separates behind the contact surface and the flow field is no longer symmetrical with respect to the symmetry line of the nozzle. The 2-D computation, on the other hand, is nonviscous and remains, therefore, symmetrical. To obtain better agreement between the 2-D computation and the experiment, a viscous boundary condition would be required in place of the inviscid symmetry boundary condition, or the experiment would have to be redesigned.
- c) The start-up process of the nozzle ranges from 2.4 ms at the lowest pressure level ($P_{41} = 4$) to 0.9 ms at the highest pressure level ($P_{41} = 188$), for the 45° nozzle.
- d) The 90° nozzle effectively forms a divergent section by establishing its flow expansion angle in response to the particular driver pressure ratio.

REFERENCES

- [1] Opalka, K. O., "Large Blast and Thermal Simulator Advanced Concept Driver Design by Computational Fluid Dynamics," BRL-TR-3026, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005, August 1989.
- [2] Opalka, K. O. and Pearson, R. J., "Real Time Flow Control in Large Blast/Thermal Simulators," Proceedings of the 1988 Army Science Conference, Ft. Monroe, Hampton, Virginia, 25-28 October 1988.
- [3] Opalka, K. O., "Large Blast-Wave Simulators (LBS) With Cold-Gas Drivers: Computational Design Studies," BRL-TR-2786, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005, March 1987.
- [4] Amann, H.-O., "Vorgaenge beim Start einer ebenen Reflexionsduese," Bericht Nr. 9/68, Ernst-Mach-Institut, Freiburg i. Br., Eckerstrasse 4, 1968.
- [5] Opalka, K. O. and Mark, A., "The BRL-Q1D Code: A Tool for the Numerical Simulation of Flows in Shock Tubes with Variable Cross-Sectional Areas," BRL-TR-2763, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005, October 1986.
- [6] Hisley, D. M., and Molvik, G. A., "Axisymmetric Calculations for the Large Blast/Thermal Simulator (LB/TS) Shock Tube Configuration," BRL-TR-2935, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, 21005, September 1988.

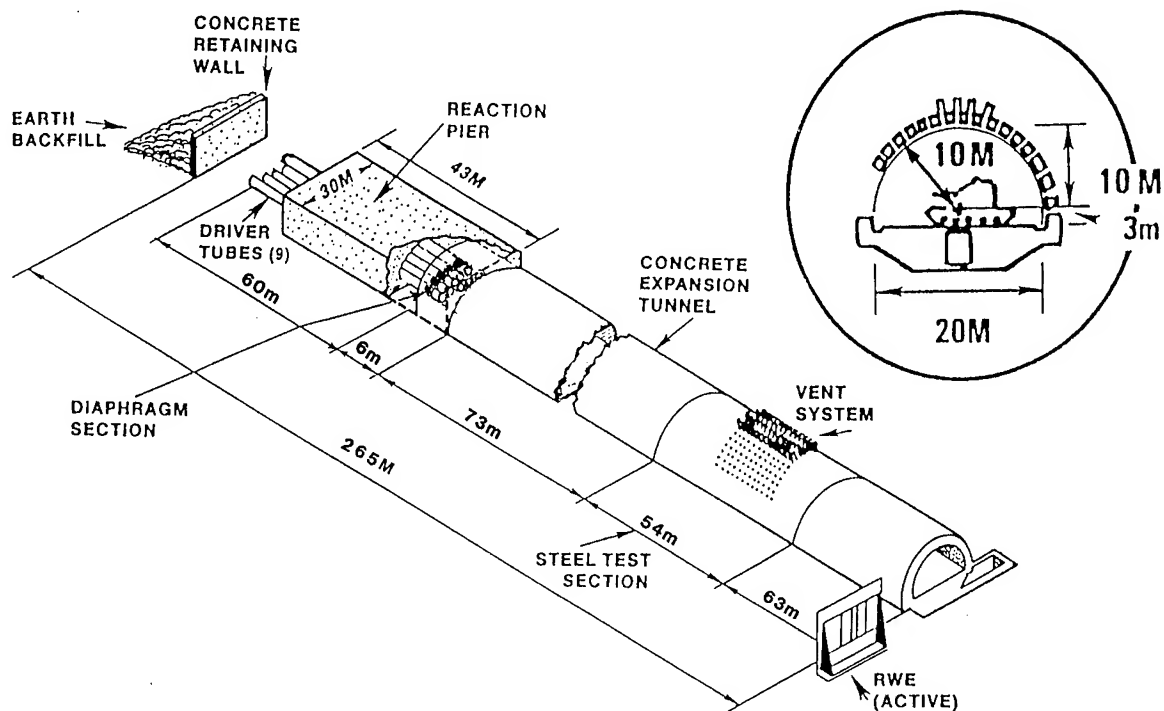


Figure 1: U.S. Large Blast and Thermal Simulator Concept

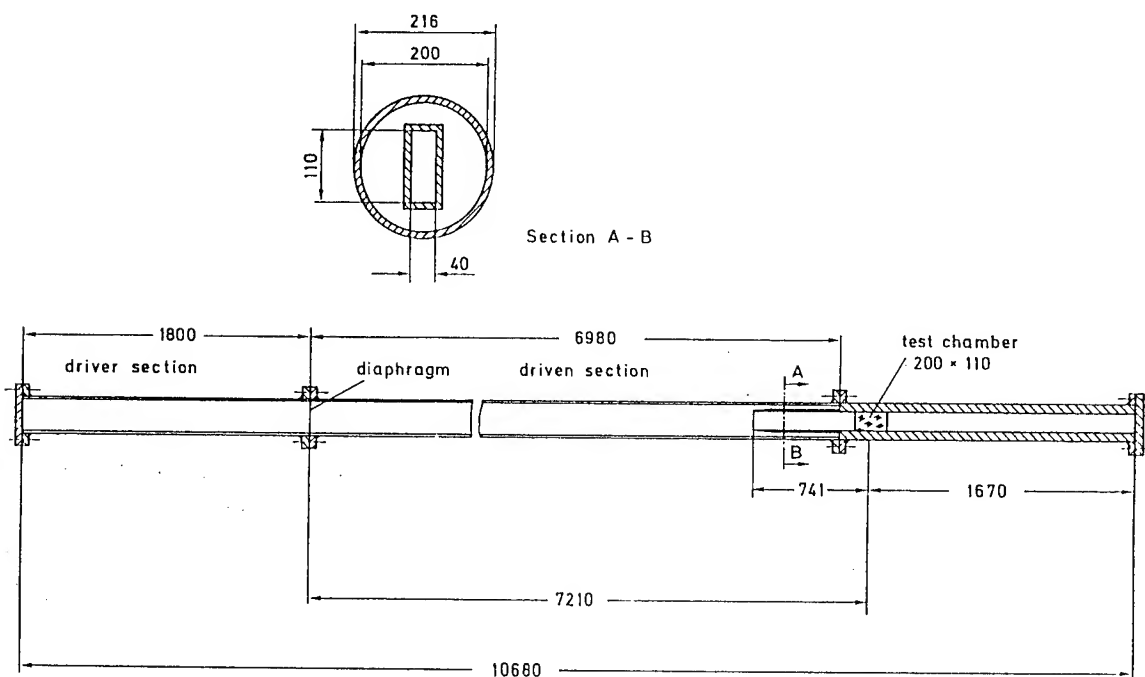
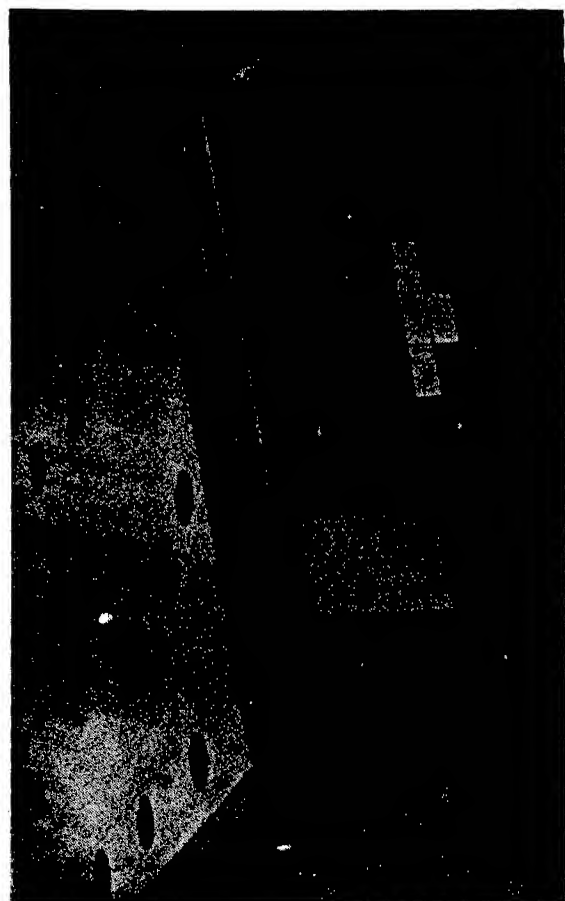
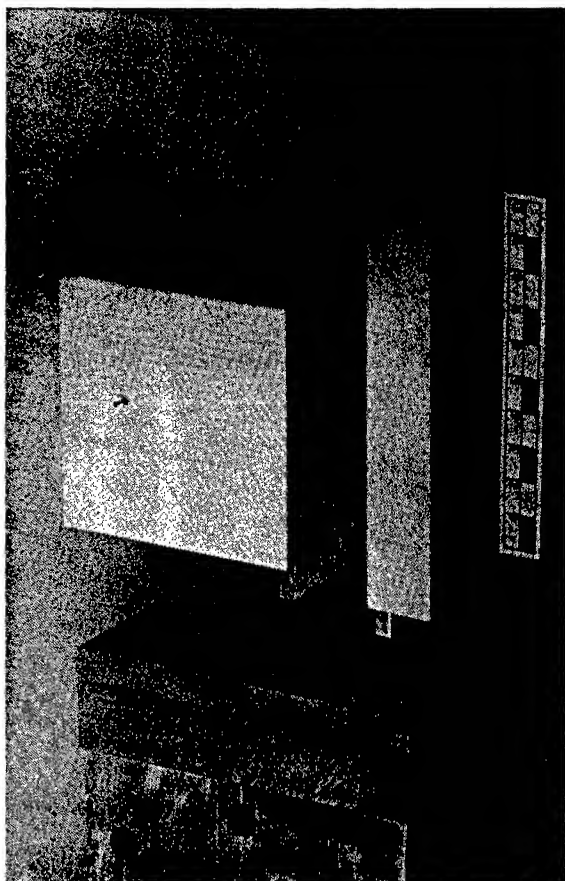


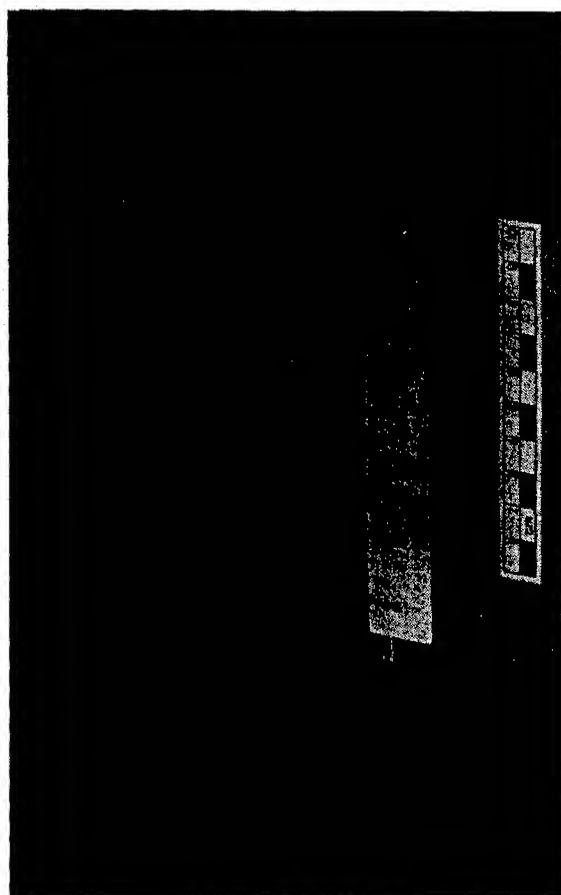
Figure 2: EMI 200mm Shock Tube (all dimensions in mm)



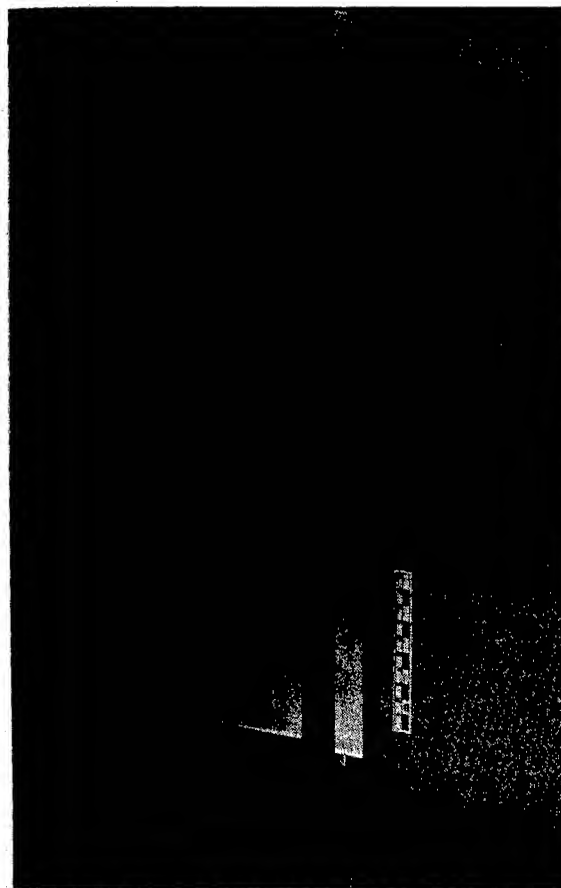
a) Diaphragm section



b) 90° nozzle

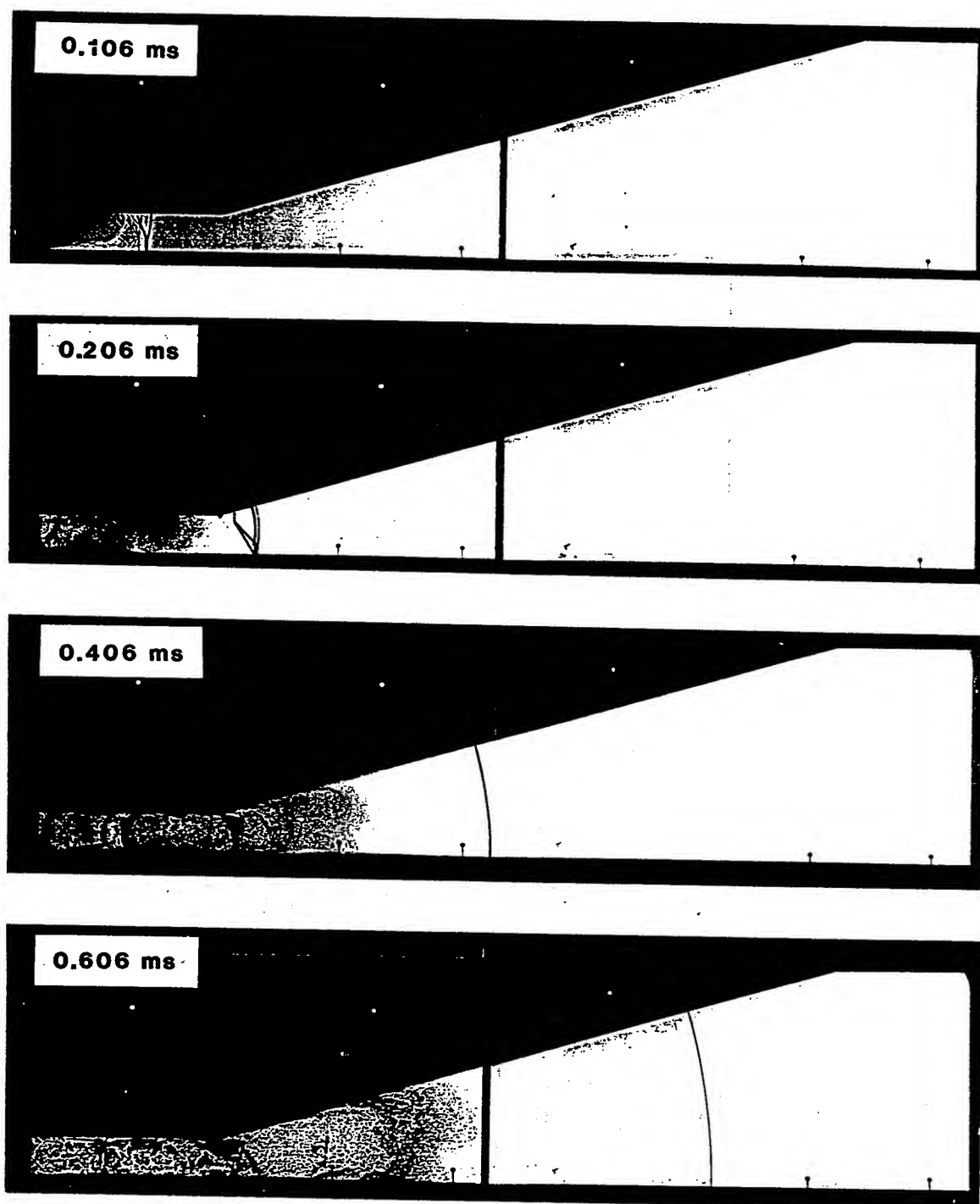


c) 450° nozzle



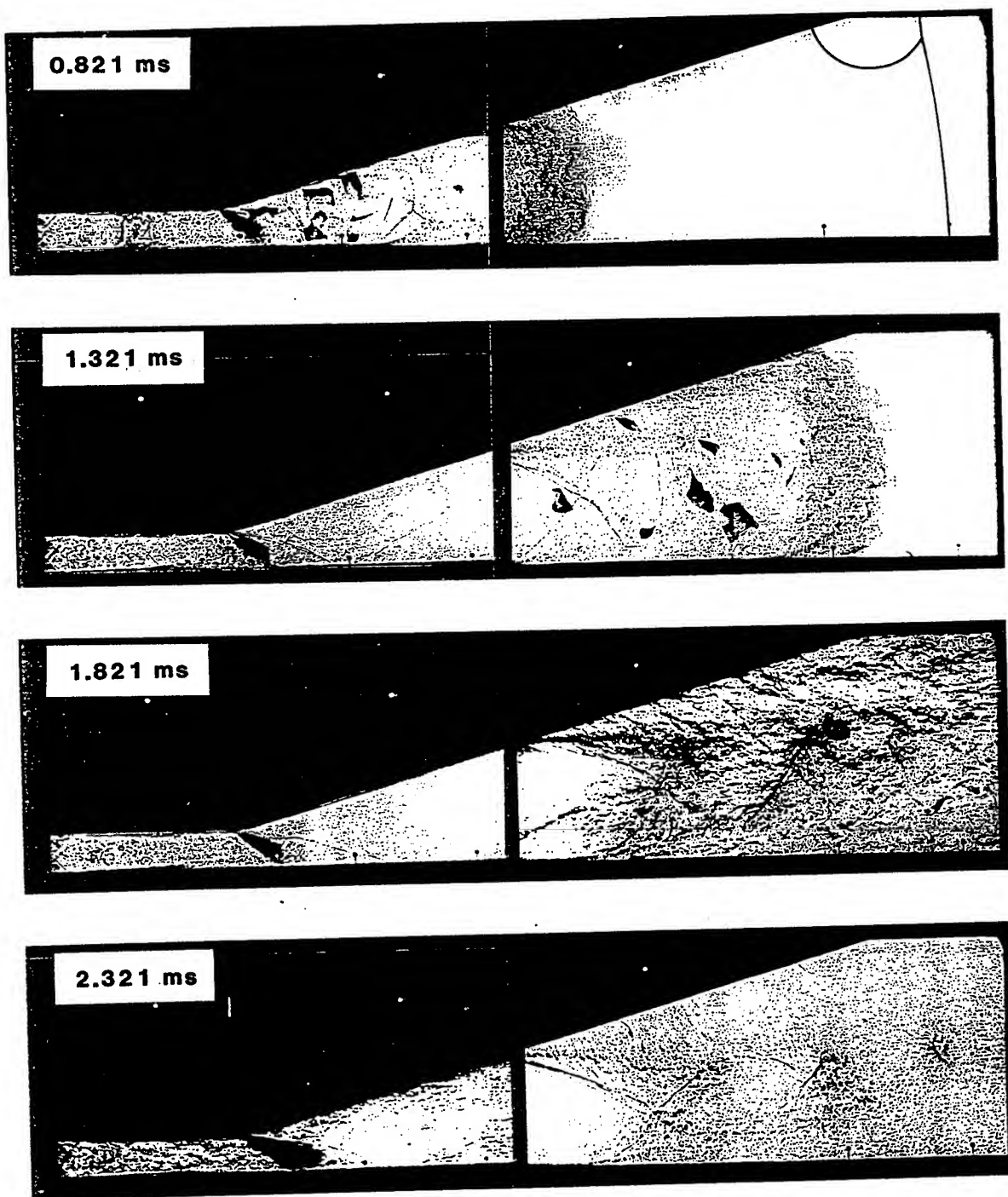
d) 6° nozzle

Figure 3: Model Driver and Nozzles



a) Early Times

Figure 4: Shadowgraph Sequence for the 16° Half Nozzle at a Driver Pressure Ratio, $P_{41} = 14$ (Code N16/014).



b) Late Times

Figure 4: (concluded)

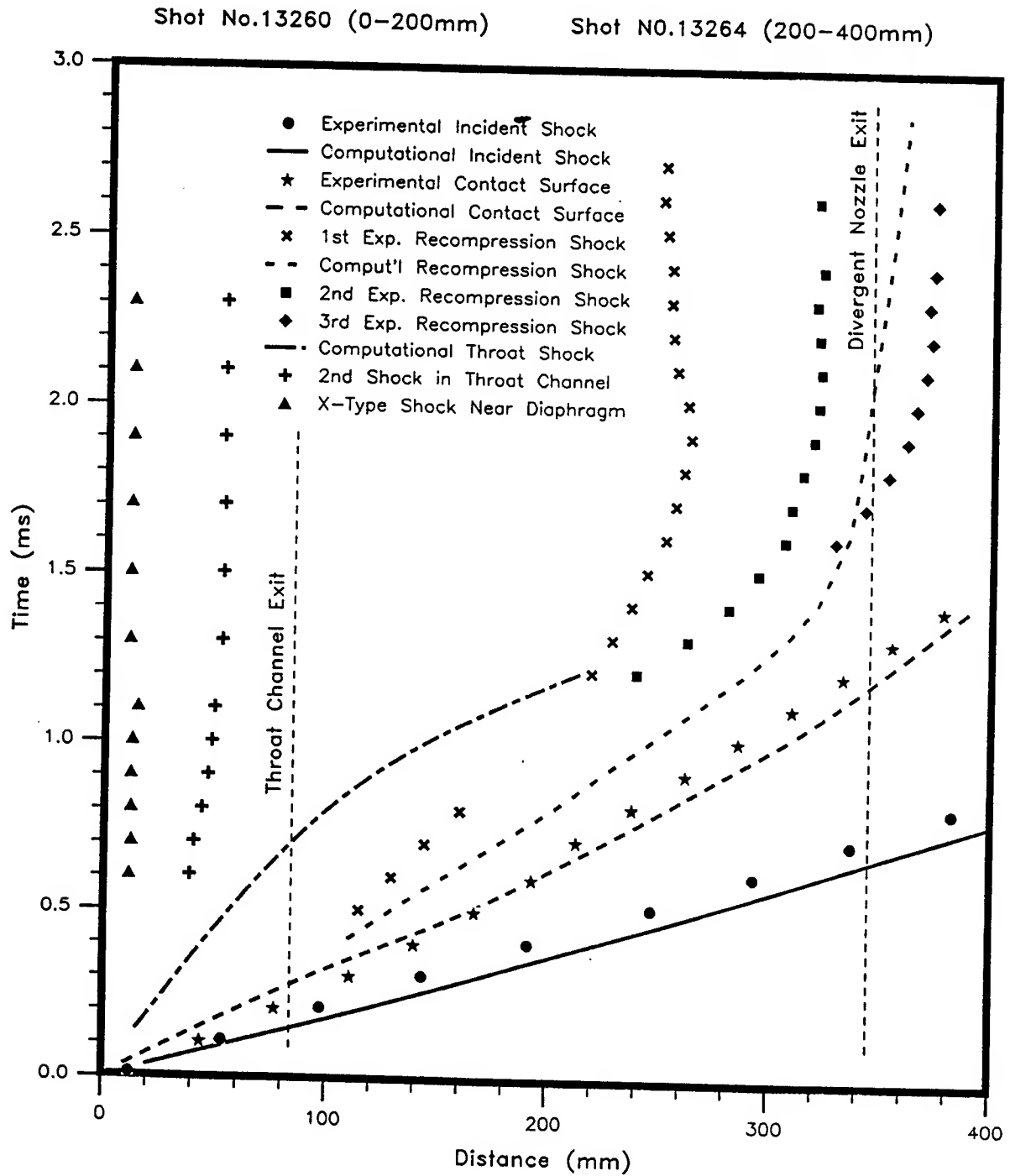
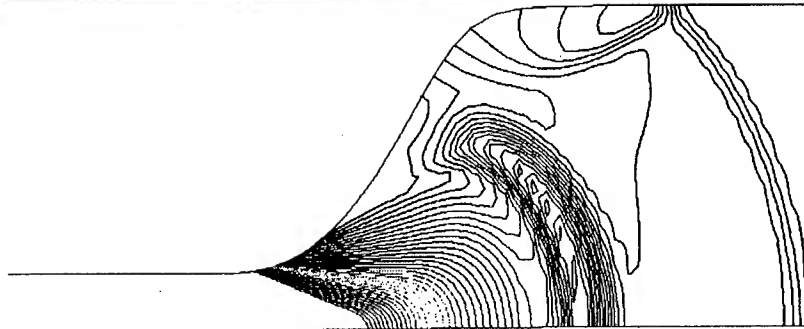
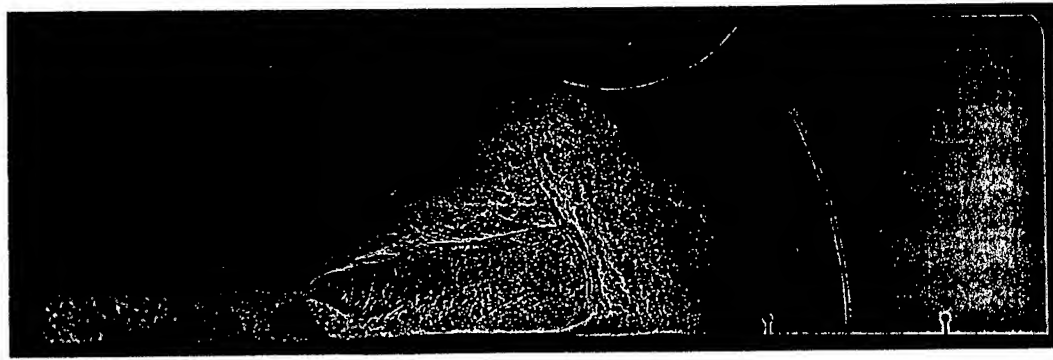
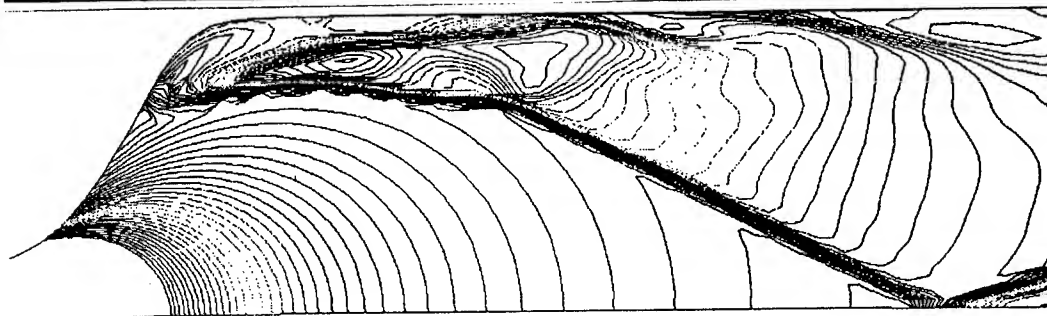
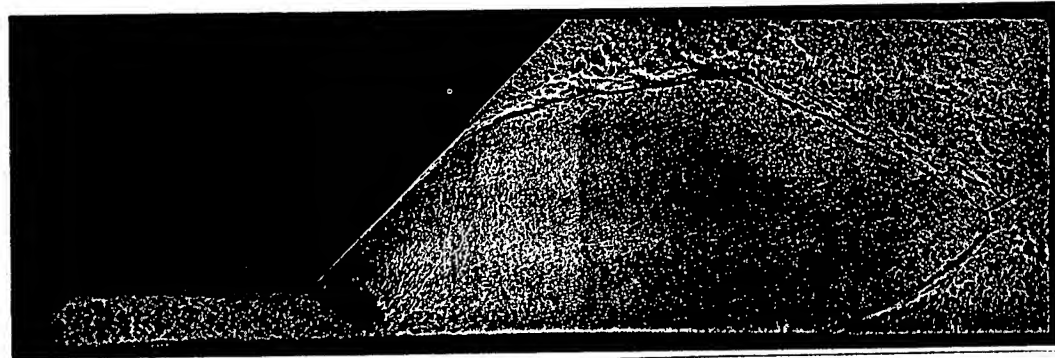


Figure 5: x-t Diagram for the 16° Half Nozzle at a Driver Pressure Ratio, $P_{41} = 14$ (Code N16/014).



a) Shadowgraph and Density Contours at 0.44 ms



b) Shadowgraph and Density Contours at 1.44 ms

Figure 6: Comparison of Experiment and Computation for the 45° Half Nozzle at a Driver-Pressure Ratio, $P_{41} = 80$ (Code N45/080)

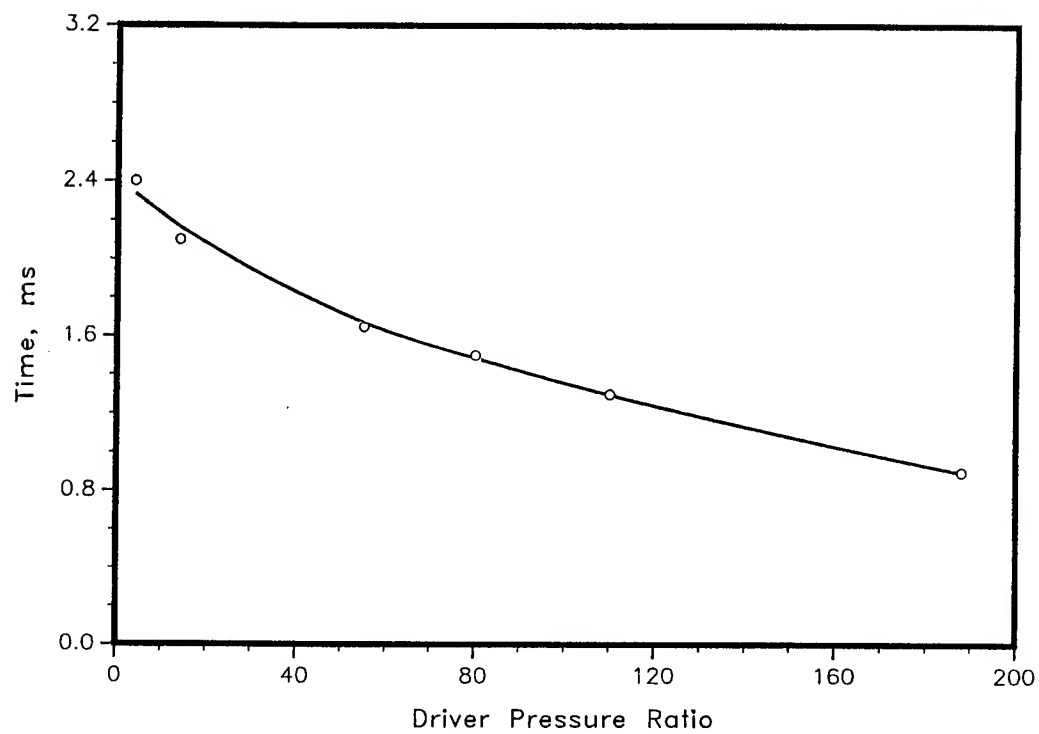


Figure 7: Flow Start-up Period for the 45° Half Nozzle

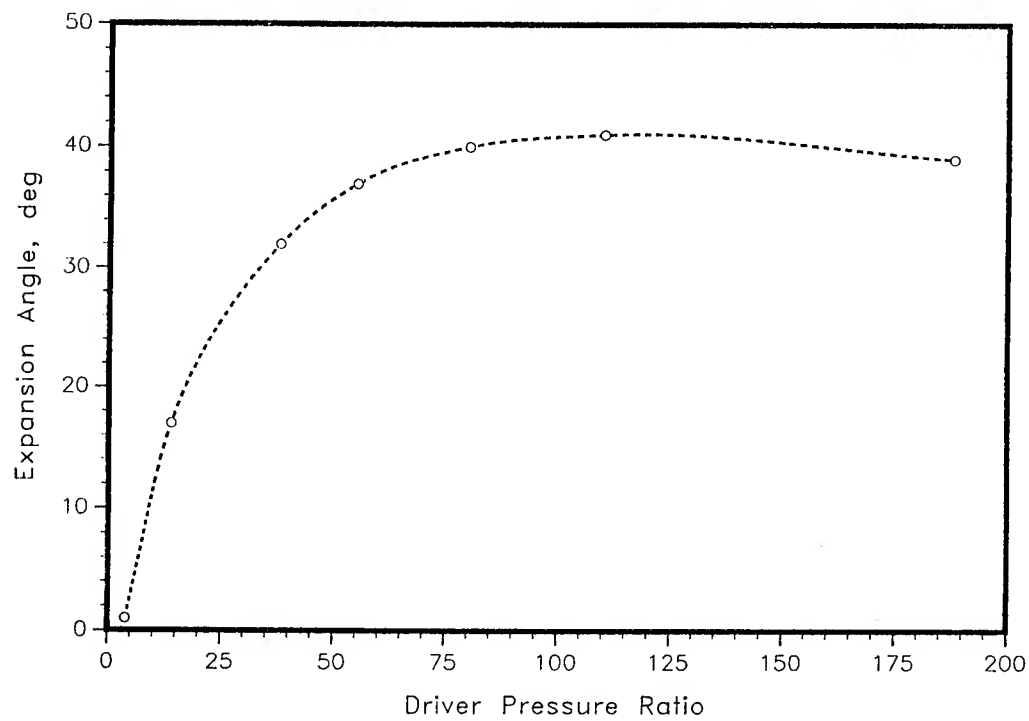


Figure 8: Flow Expansion Angle behind 90° Half Nozzle

PAOLELLA

Lightwave Technology in Microwave Systems (U)

Arthur Paolella

U. S. Army, LABCOM,
Electronics Technology and Devices Laboratory,
Ft. Monmouth, N. J., 07703

INTRODUCTION

Two complimentary technologies that have emerged in the 1980's, GaAs Microwave Monolithic Integrated Circuits (MMIC) and fiber optic links, will have a significant impact on the next generation of microwave systems. GaAs MMICs will not only reduce the size and cost of microwave systems but, combined with fiber optics, it will open the way for a wide range of applications. The technology will make systems such as active phased array radars a viable technique for implementation.

Admittedly active phased array radars use a large number of individually controlled transmit-receive (T/R) modules which require independent microwave reference and control signals. The resulting distribution networks needed to route this information to the T/R modules to perform beam steering and beam shaping can become the limiting factor in these systems. Conventional distribution networks such as coaxial transmission lines and metallic waveguide are bulky, and susceptible to radiation in hostile environment. Also conventional control lines that feed microwave circuits can also carry extraneous signals and cause errors in circuits performance. Isolation between various circuits in a system can be accomplished by using optical control signals. Lightwave transmission systems such as fiber optics have become an attractive alternative for the distribution of microwave signals and control signals due to two factors; the low cost of fibers, and the increasing speed of lasers and photodetectors. Distributing the various control signals by optical fibers have desirable features, such as high speed, large bandwidth, good electrical isolation, and elimination of grounding problems. Also, the fiber is lightweight, has a small size, and most significantly, it is immune to electromagnetic interference (EMI) and electromagnetic pulses (EMP). The size of the fibers themselves and the electrooptic components such as semiconductor lasers and photodetectors make them highly compatible with GaAs MMICs. In addition, the inclusion of optical techniques for the control of microwave subsystems, provides an opportunity for the introduction of optical computing and signal processing techniques for beam steering and beam forming.

PAOLELLA

TRANSMISSION OF MICROWAVE SIGNALS

Each T/R module in an active phased array antenna is an independent entity. This requires the synchronization of each module to a reference signal, usually a microwave carrier frequency. Several architectures have been investigated to determine individual capabilities to distribute this high frequency signal via a fiber optic link. These architectures are the object of the following discussion.

The basic link, as shown in Fig. 1, consists of a high speed semiconductor laser, a low loss fiber optic cable, a high speed photodiode, and microwave matching circuitry. It is preferable to use semiconductor electrooptic components based in GaAs as to maintain compatibility with GaAs MMICs in anticipation of the integration of electrooptic and microwave components on the same chip.

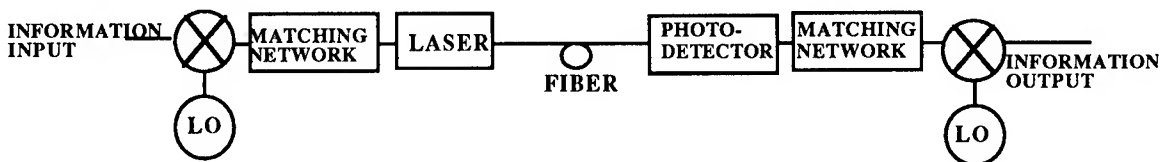


Figure 1. Block diagram of a fiber optic link used to transmit microwave signals.

Presently semiconductor lasers which operate at a wavelength of 850 nm are commercially available and can be modulated to 10 GHz. These lasers are buried heterojunction GaAlAs structures. The maximum modulation frequency of this type of laser is determined by the relaxation oscillation frequency. This frequency in turn is determined by many factors, including device structure, photon density, and photon lifetime. Above the relaxation oscillation frequency, the attenuation of the microwave signal becomes significant, and the relative intensity noise (RIN) rises, limiting the laser's use in the fundamental mode of operation. The optical fibers available for use in these links are numerous with losses less than 1 dB/Km. A typical size fiber is between 5 μ m to 125 μ m core diameters. The photodetectors used for these links are PIN device structures which can be used to 20 GHz. With these aforementioned components, microwave signals along with modulation information, can be transmitted via fiber optic links by direct modulation of the laser well into X-band, which covers many types of radar applications.

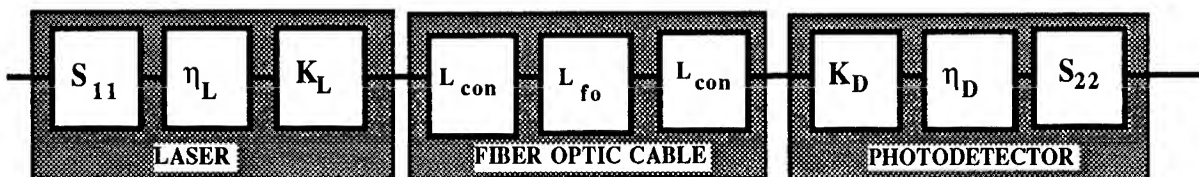


Figure 2. Losses in an analog fiber optic link.

PAOLELLA

The efficient transmission and detection of microwave signals via fiber optic links requires that several loss mechanisms, as depicted in Fig. 2, be reduced. The gain (or loss) of a link can be calculated by the following equation:

$$G[\text{dB}] = 10 \log(L_{\text{mw}}^2 L_{\text{rsp}}^2 L_{\text{co}}^2 L_{\text{fo}}^2 L_{\text{con}}^2)$$

where:

$$L_{\text{mw}}^2 = (1 - |s_{11}|^2) (1 - |s_{22}|^2) \frac{R_D}{R_L}$$

S_{11} , S_{22} are the reflection coefficients of the laser and detector

R_L , R_D are the laser and photodetector resistances

$$L_{\text{rsp}}^2 = \eta_L^2 \eta_D^2$$

η_L , η_D are the responsivities of the laser and detector

$$L_{\text{co}}^2 = K_L^2 K_D^2$$

K_L , K_D are the laser and detector coupling efficiency and

$$L_{\text{fo}}^2 L_{\text{con}}^2$$

are the fiber losses and connector losses.

The design of the link requires careful optimization of the microwave circuitry as well as of the optical components. It may be readily seen in the above equation that the losses from the reflection coefficients S_{11} and S_{22} must be minimized over the desired bandwidth with respect to a typical 50 ohm system. Typical losses are in the order of 13 dB [1]. Poor input matching results from the low impedance of the laser which is typically 3 ohms with a capacitive reactance in parallel. Poor output match results from the high impedance of the reversed biased PIN photodetector which is mostly capacitive. The reflections from the laser and detector can be reduced through the design of passive matching circuits between the microwave system and the laser, and between the photodetector and the microwave output. Matching networks can be fabricated using microstrip transmission lines or lumped element circuits. A second high loss mechanism is the coupling of light from the laser to the fiber optic cable at the transmitter, and the coupling of the optical signal from the fiber to the photodetector. The losses in these short haul systems can be reduced by using multimode fiber with large core diameters, and by careful alignment of the fiber to the electrooptic component.

The dynamic range of the fiber optic links is limited by the relative intensity noise of the laser. Dynamic range of 22 dB at 6 GHz has been reported [1]. To improve the dynamic range

PAOLELLA

requires either redesigning the structure of the electrooptic components, or by using other link configurations. It has been found by I. Kauffman, et al., that by transmitting the microwave carrier over a high speed, narrow band FO link, and sending the low frequency modulation information (typically 500 MHz to 1 GHz) via a wideband low frequency link, and mixing the signal at the module, improvements in dynamic range of 20 dB over conventional links has been achieved. This type of remote coding has been demonstrated with encouraging results. To date the best performance of an analog optical link was reported by Ackerman, et. al.. Their directly modulated optical data link produced 3.7 dB gain at 900 MHz over a 90 MHz bandwidth, with dynamic range of 89 db. These results were achieved by using passive matching techniques only[2].

Above the relaxation oscillation frequency of the laser alternate techniques are needed to send signals in the sub-millimeter wave and millimeter wave regions. For this purpose, advantage can be taken of the non-linearities of the laser and the microwave circuitry to injection lock a millimeter wave oscillator [3]. In this configuration the laser is modulated at a sub-harmonic of the carrier frequency. After all the harmonics are transmitted optically to a photodetector, the desired harmonic is amplified and used to sub-harmonically injection lock a millimeter wave oscillator. It has been reported by Daryoush, et. al., that a 38.820 GHz oscillator was injection locked by a laser which was modulated from a 12th sub-harmonic(3.235 GHz). An injection locked bandwidth of 132 MHz was then obtained.

It should be noted that improvements in laser technology such as growth of materials and processing have allowed system designers to use semiconductor lasers operating at 1.3 μm . These lasers by their material properties can operate as high as 15 GHz.

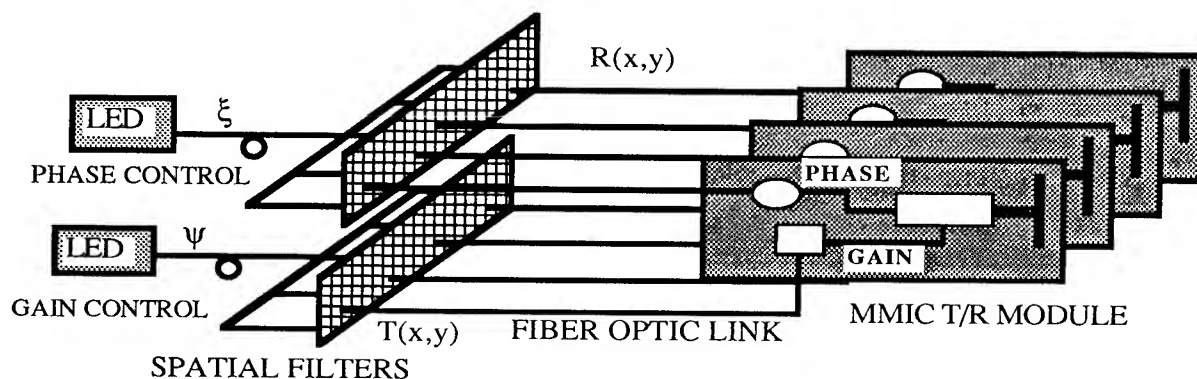


Figure 3. Optical control of T/R modules using spatial filters.

PAOLELLA

OPTICAL CONTROL OF MICROWAVE CIRCUITS

In microwave systems it is preferred that components and sub-systems be isolated so that radiation from other systems do not interfere with each other. Conventional control lines that feed microwave circuits such as switches, phase shifters, etc., can carry extraneous signals and cause errors in circuit performance. Isolation between various circuits in a system can be accomplished by using optical control signals. Fiber optic cables can carry information to various circuits or sub-systems without picking up extraneous signals such as EMI and EMP. Fiber optic cables can also be used for the distribution of high density signals such as gain and phase control in active phased array radars. Examples of the uses of fiber optics in the control of microwave circuits will follow along a presentation of the obtained with results.

For a one dimensional phased array with N nondirectional elements shown in Fig. 3, the field strength in the far field is given by

$$E_T = E_O \sum_{n=0}^{N-1} K_n e^{j\phi_n}$$

where K_n is excitation to the n^{th} element, and ϕ is the phase shift. In an active array antenna the excitation to each individual element is controlled, allowing for amplitude tapering and beam steering. In a T/R module, the excitation to each antenna element is a function of the amplifier gain and phase, thus

$$K_n = K_O G_n \text{ and } \phi_n = \phi_O + F_n$$

where K_O is the input to each amplifier in the array, G_n is the amplifier gain of the n^{th} element, and ϕ_O is an initial phase to each phase shifter in the array. F_n is the additional phase shift of the n^{th} element required to steer the beam in the desired direction. For optically controlled beam forming, the gain and phase of each module must be a function of optical intensity such that

$$G_n = G_n(\psi_n) \text{ and } F_n = F_n(\xi_n)$$

where ψ_n is the optical intensity to the n^{th} amplifier, and ξ_n is the optical intensity to the n^{th} phase shifter.

If in fact efficient optical gain and phase controllers can be developed, then the antenna beam pattern can be controlled by spatial light modulators as in Fig. 3. For spatial light modulators with a transfer function $T(x,y)$ and $R(x,y)$

$$\psi_n = T(x,y)\psi_O \text{ and } \xi_n = R(x,y)\xi_O$$

PAOLELLA

where ψ_0 and ξ_0 are the uniform input optical power illuminating the spatial light modulators to control the gain and phase respectively. Our effort was directed towards the development of a simple low-cost technique which would be compatible with the state of the art MMIC processing and fabrication methods, and capable of controlling the gain and phase of a T/R module optically, thereby producing $G_n(\psi_n)$ and $F_n(\xi_n)$. Ideally one would like to interface the optical input with the MMICs directly. However PIN photodiodes, commonly used with fiber optic systems, are not compatible with standard MMIC processing schemes for compatibility with GaAs MMIC technology a MESFET was chosen as an optical detector.

There are three basic modes of operation of the MESFET when used in conjunction with fiber optics: **active**, **pinch-off** and **avalanche**, each determined by the bias condition and by the control approach; i. e., **direct or indirect**. In the **active** mode the MESFET is used as an amplifier or oscillator as it would normally be in a microwave circuit. Here the device is biased around 50 % to 100 % of I_{DSS} and the gate-to-source voltage is low (0 to -2.0 volts). The optical signal is **directly** illuminating the device, changing its operating conditions. This approach is employed to control amplifier gain or for injection locking of an oscillator. In the **pinch-off** mode the MESFET is used as an optical detector to receive a control signal. Here the device is in a common source configuration with a resistor from the bias supply to the drain of the MESFET as in a voltage divider circuit. When the device is illuminated, the optically generated drain current changes the drain-to-source voltage of the MESFET, which is sampled, amplified, and then utilized to control the biasing of an amplifier or phase shifter for gain and phase control respectively. In this mode, the device is biased with a large gate-to-source voltage, around pinch-off (-3.0 to -4.0 volts) where the optically generated drain current produces the maximum voltage change. This approach is referred to as **indirect** optical control since the optical signal is first detected and then an electrical control signal is applied to the control element. In the **avalanche** mode the device is biased well beyond pinch-off with a gate-to-source voltage set around -4.5 to -6.0 volts (drain current <100 μ a), and with large drain-to-source voltages (6.0 to 8.0 volts) such that photo-multiplication effects are introduced[4]. In this mode, large optically induced drain current is achieved even at low level illumination, on the order of a few microwatts, due to avalanching effects in the depleted region (i.e., between the gate and drain).

To produce gain changes optically, previous investigators [5-7] have concentrated on the control of high speed MESFET by injecting light into the active region between the gate and the source, or the drain. These experiments were limited by the poor coupling of the light into the active region of the device. The long and narrow geometry of high speed, high frequency MESFET transistor is fundamentally incompatible with a typical cylindrical gaussian light spot emerging from a fiber. Improvements in gain and phase changes by optical means have been developed by controlling the biasing circuit of an amplifier and the control line of a variable bit phase shifter. The optical sensing element was a multi-finger MESFET which was illuminated by an LED source. The optical injection effectively controlled the gate to source voltage of the MESFET which in turn was used to control the bias voltage of the amplifier and the control voltage of the phase shifter and hence the gain and phase of the module.

PAOLELLA

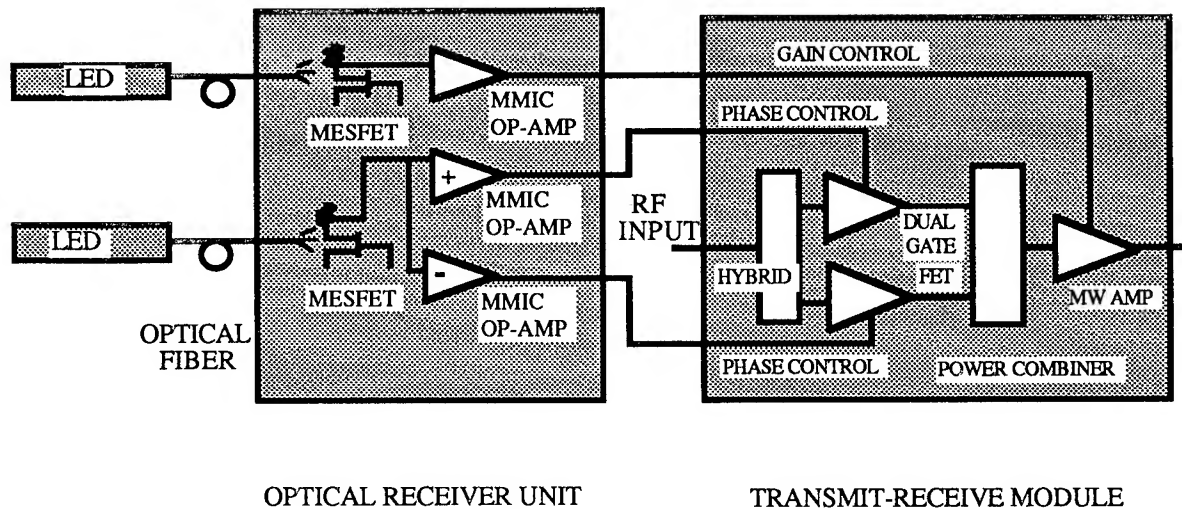


Figure 4. Optical gain and phase control of a GaAs MMIC T/R module.

The experimental setup is established based on the block diagram shown in Fig. 4. However, the actual experimentation was carried out separately for the gain and phase control because the MMIC phase shifter and distributed amplifier were obtained separately. The gain control element was an MMIC distributed amplifier whose performance was controllable through its bias voltage [8,9]. The phase shifter used for the experiments consists of a combination of fixed bit phase shifters along with a single quadrature vector modulator [10]. For a 45 degree phase shift less than 8 dB of gain control is required; a higher degree of phase shift can be obtained by cascading several sections of the basic vector modulator.

There are four main elements to the system: i) a pig-tailed LED source, ii) a light sensitive MESFET used as a photodetector, iii) a d.c. operational amplifier, and iv) the T/R module containing the phase shifter and amplifier. All of these components, except the LED, are GaAs MMICs. The optical source is an LED which provides a maximum of 300 μW of optical power centered at 835 nm. The critical part of the arrangement is the coupling of the light to the MMIC chip. In these experiments a 14 gate finger MESFET served as an optical detector. By using a multi-finger MESFET the exposed GaAs area for light absorption was increased resulting in a significant improvement in the optical response.

Referring to Fig. 5, the optical gain control of the distributed amplifier was accomplished by the following method. With the MESFET biased near pinch-off, the drain-to-source voltage was sampled by an MMIC operational amplifier operating in the inverting mode with 10 dB of gain. The output of the op-amp was connected to the three gates of the distributed amplifier providing the bias needed to control the gain. The gates of the distributed amplifier were biased at pinch-off by an offset voltage from the op-amp so that the amplifier gain was low. As the optical illumination incident on the MESFET was increased, the gain of the distributed amplifier increased. A change in $|S_{21}|$ of -10 dB to +5 dB was achieved over the frequency range of 5 to 8 GHz with only 250 μW of optical power, as shown in Fig. 6.

PAOLELLA

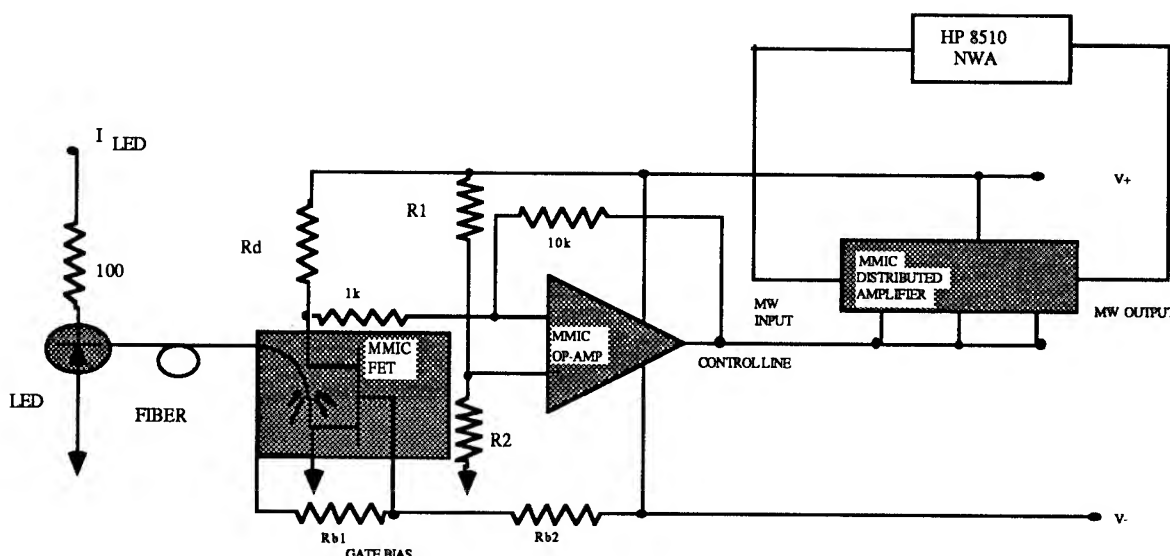


Figure 5. Experimental set up for the optical gain control of a microwave distributed amplifier. The experiment consists of i) an LED with a pig-tailed optical fiber ii) a GaAs MESFET as an optical detector, iii) a GaAs MMIC op-amp, iv) a GaAs MMIC distributed amplifier.

Measurements of gain as a function of optical intensity were made on a network analyzer. The experiment for the optical phase control was an extension of the optical gain control technique used in the previous experiment[8]. A 360 degree L-band continuously variable phase shifter (CVPS) was controlled over a 45 degree sector (-22.5 to +22.5) by varying the optical intensity of one LED. The phase shifter, which is implemented as a vector modulator, consists of three fixed bits to select a given 45 degree sector of operation, two quadrature amplifiers to produce continuous phase shift, and a final gain stage to control the overall phase shifter gain. In this configuration, the optical intensity of a single LED is varied to produce the appropriate control voltages for the quadrature amplifiers. The gain of the amplifier is controlled by the bias voltage, which in this configuration depends on the optical signal to the MESFET detector. Phase shift is obtained by changing the bias voltages on the dual-gate MESFET amplifiers, which operate in quadrature. The output is then combined to give the proper phase to the optically controlled variable gain amplifier.

Referring back to Fig. 4, the phase shifter operates in the following manner. One op-amp is in the inverting configuration with 10 dB of gain. The other op-amp is in the non-inverting mode. Varying the light intensity gives rise to a change in the drain-to-source voltage (V_{gs}) which, after amplification, provides a necessary control voltage for the phase shifter. The plot of phase shift vs. optical intensity over the phase shifter bandwidth is given in Fig. 7, demonstrating a phase shift as much as 45 degrees for $51\mu W$ of optical power. The maximum variation in the peak amplitude over the full 45 degree sector is 1.6 dB.

PAOLELLA

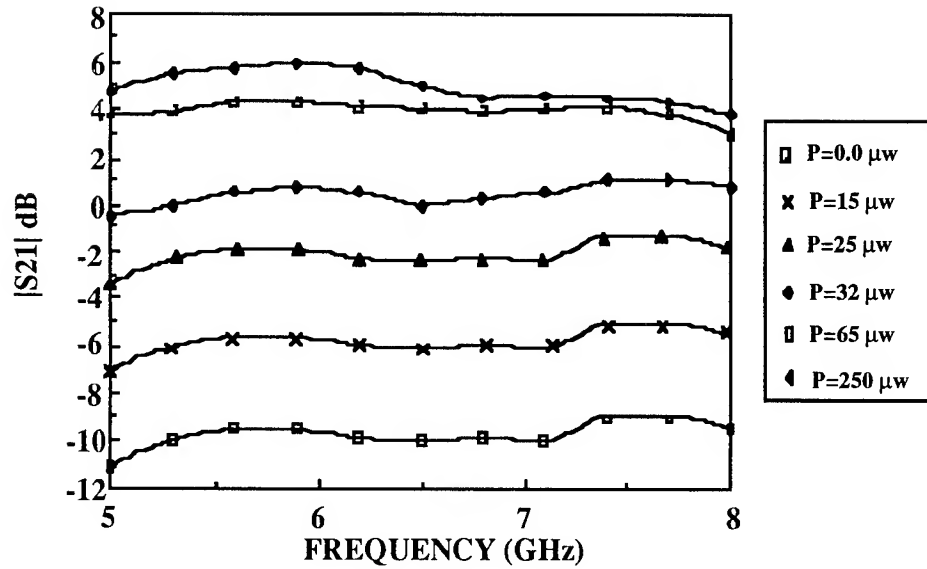


Figure 6. Experimental results of $|S_{21}|$ of the microwave distributed amplifier as a function of optical intensity.

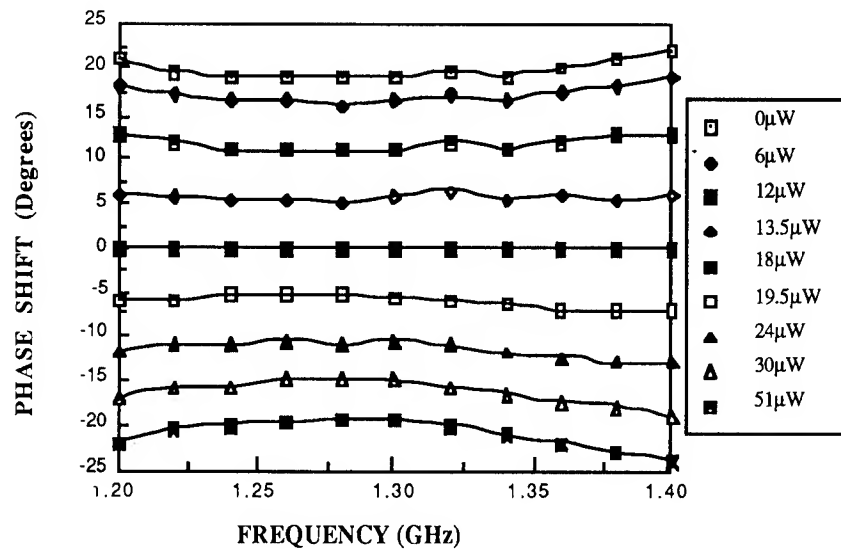


Figure 7. Experimental results of the phase of the vector modulator phase shifter as a function of optical intensity.

PAOLELLA

Improvements in phase control have been accomplished. In another experiment, optical control of a 6-bit X-band phase shifter where an A/D converter is used to generate the required parallel digital phase shifter command has been demonstrated[11]. Only one fiber optic input for each phase shifter is required in addition to an A/D encode command. This approach is applicable to all known digital phase shifters, regardless of their operating frequency. An all digital phase shifters consisting of n-bits may be optically controlled by an intensity modulated LED or laser diode. Again a MESFET is used as the optical detector. The intensity is varied to produce 2n equally spaced discrete MESFET output voltages, which are appropriately scaled via a standard amplifier to correspond to an A/D converter input voltage range. The A/D converter then converts the voltages to an n-bit binary word which is used to command the phase shifter to a desired phase state. In this way, the intensity level of the incident optical input sets the phase shifter to the desired state. A block diagram of the phase shifter is shown in Fig. 8 and the test results at 10 GHz are shown in Fig. 9.

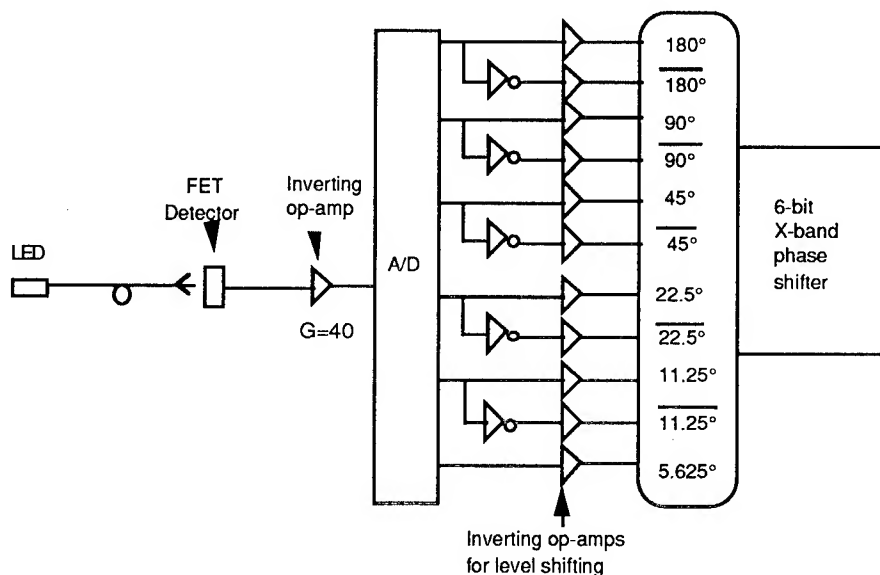


Figure 8. Block diagram of an optically controlled 6-bit X-band phase shifter where an A/D converter is used to generate the required parallel digital phase shifter command.

Another method to obtain phase shift is to use true time delay in the optical domain. In a fiber optic link the phase difference from input of the link to the output of the link is determined by the distance the signal travels in the fiber and the propagation velocity as given in the following equation

$$\Delta\theta = 2\pi f \tau_t = 2\pi f L v_g$$

PAOLELLA

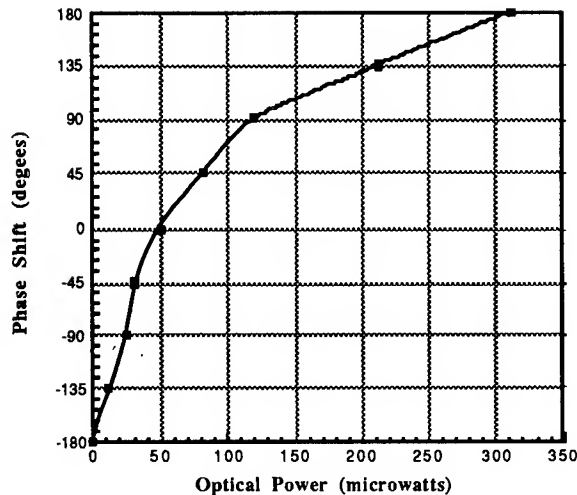


Figure 9. Phase shift as a function of optical power for an X-band digital phase shifter.

True time delay phase shift can be obtained by switching in different line lengths. Variable phase shifts can be obtained using a technique whereby the fiber is rapped around a PZT and is stretched when an electric field is applied across the PZT [12]. The results of this experiment showed a 20 degree shift in phase at 10 GHz with good repeatability, linearity, and no hysteresis. This technique, however, is limited in speed and requires high operating voltage. New materials be tested with a view at improving performance.

Other optically triggered circuit functions have been demonstrated. An optically controlled GaAs MMIC switch was constructed and tested again using the MESFET as an optical [13]. The experimental setup is shown in Fig.10. The microwave switch, contained on a single GaAs MMIC chip, is a single pole double throw (SPDT) which operates from 8 to 12 GHz. The switch, manufactured by MA/COM, contains 5 MESFETs which require no biasing. The switching function is controlled by two voltage settings, V_1 and V_2 . To switch the microwave input signal to OUTPUT 1, the required voltages are $V_1=0.0$ and $V_2=-7.0$ volts. To switch the microwave input signal to OUTPUT 2, the voltages needed are $V_1=-7.0$ and $V_2=0.0$ volts. The control circuitry is fabricated on 25 mil alumina using laser trimmed thick film resistors, unpackaged GaAs op-amps, and an MMIC chip MESFET for the detector. The GaAs MMIC switch was mounted on a separate 25 mil alumina substrate which contained lines for the control voltages, and three 50 ohm microstrip transmission lines for the microwave signals. the circuit size measures approximately 1 inch squared.

The operation of this circuit follows. The MESFET optical detector is biased at pinch-off. The biasing to the gate circuit is provided by resistors R_{G1} and R_{G2} . The ratio is set to provide a gate-to-source voltage (V_{GS}) of -3.75 volts. The biasing to the drain circuit is provided by the

PAOLELLA

resistors R_{D1} and R_{D2} . Here the ratio is set so that at pinch-off the drain-to-source voltage (V_{DS}) is 3.0 volts. Under optical illumination the device begins to conduct current through R_D and the required change in V_{DS} of 0.5 volts is obtained.

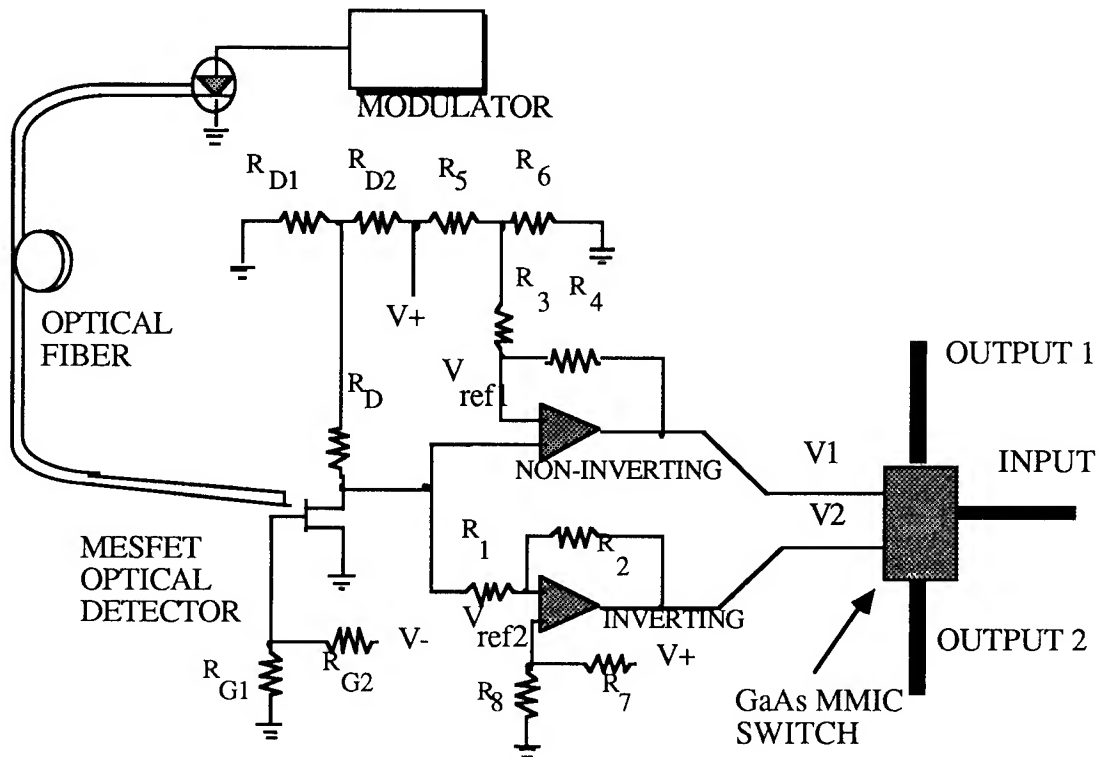


Figure 10. Optically controlled GaAs MMIC switch.

The optical source is a fiber coupled LED operating at a peak wavelength of 835 nm. The light is routed via a multi-mode fiber with core-cladding diameters of 100 and 140 μm respectively. The optical intensity needed to provide the 0.5 volt change is only 25 mW. The drain of the MESFET is connected to two high speed operational amplifiers with a voltage gain of 14 (23 dB) and a gain-bandwidth product of 400 MHz. One of the amplifiers operates in the inverting and the other in the non-inverting mode. The non-inverting op-amp has a reference voltage, $V_{ref1}=3.0$ supplied by voltage divider resistors R_5 and R_6 , while the resistors R_3 and R_4 provide voltage gain. In the absence of illumination, $V_{DS}=3.0$ volts, the difference in the input voltage of the non-inverting op-amp is 0.0 volts, and therefore the output voltage, V_1 , is 0.0 volts. The inverting op-amp has a reference voltage of $V_{ref2}=2.5$ volts supplied by resistors R_7 and R_8 . Gain for this op-amp is set by R_1 and R_2 . Again with $V_{DS}=3.0$ volts, a difference of 0.5 volts exists at the input of the inverting op-amp. The gain is such that the

PAOLELLA

output V_2 is -7.0 volts. With these conditions, OUTPUT 1 is in the low loss state, and OUTPUT 2 is in the isolation state.

When the MESFET is illuminated, V_{DS} changes from 3.0 to 2.5 volts and the outputs of the op-amps switch states. The voltage difference at the inverting op-amp input is now 0.0 volts, while the voltage difference at non-inverting op-amp input is -0.5. The voltage V_1 swings from 0.0 to -7.0 volts and the voltage V_2 changes from -7.0 to 0.0 volts, thereby switching the microwave signal from OUTPUT 1 to OUTPUT 2. Switching rates of 100 ns have been obtained.

CONCLUSION

There is increasing interest in using fiber optics in microwave systems for distributing microwave signals, modulation information, control signals, and for signal processing. Of particular interest is the use of optics in GaAs based MMIC active phased array antennas. The distribution of microwave and millimeter wave signals has been demonstrated here and several configurations were investigated which have improved dynamic range. Also the optical control of the gain and phase of T/R modules has also been demonstrated, as well as switching functions of microwave MMICs. These applications can be used to steer and shape the radiation pattern of active phased array antennas. Time delay signal through fibers was also developed. Although the optimization of these circuits requires further investigation, the initial results are encouraging. In the future, several research issues must be addressed. They include higher speed electrooptic components, thorough system analysis and system integration. In the area of MMICs, it is foreseen that future work will address the integration of electrooptic and microwave components on the same chip.

It should be further noted that fiber optics in microwave systems is not limited to radar. Other applications have been demonstrated for use in memory loops[14] and in optical tuning of millimeter wave oscillators[15].

ACKNOWLEDGEMENTS

The author would like to thank the Center for Microwave and Lightwave Engineering, Drexel University, Phila., Pa., and in particular Dr. P. R. Herczfeld, and Dr. A. S. Daryoush for their valuable support, and to Mrs. Miriam Paolella for her editing of this paper.

REFERENCES

- [1] Kauffman, I., Herczfeld, P. R., Daryoush, A. S., "High Speed Fiber Optic Links for Short Haul Microwave Applications", IEEE MTT-S International Microwave Symposium, New York, New York, May, 25-27, 1988.
- [2] Akerman, E., Kasemet, D., Wanuga, S. Hogue, D., Komiak, J. "A High-Gain Directly Modulated L-Band Optical Link", 1990 IEEE MTT-S International Microwave Symposium, Dallas Texas, May 1990.

PAOLELLA

- [3] Daryoush, A. S., et al., "Indirect Sub-harmonic Optical Injection Locking of a Millimeter Wave IMPATT Oscillator", IEEE Transactions on MTT, Vol. MTT-34, Num. 12, Dec. 1896.
- [4] Madjar, A., Paoella, A., Herczfled, P. R., "Photo-Multiplication Effects in GaAs MESFETs., *Microwave and Optical Technology Letters*, Vol. 2, No.11, February 1990.
- [5] Simons, R. and Bhasin, K. B., "Microwave Performance of an Optically Controlled AlGaAs/GaAs High Electron Mobility Transistor and GaAs MESFET", Proc., IEEE MTT-S International Microwave Symposium, Las Vegas, Na.,1987.
- [6] Stabile, P., Rosen, A., Herczfled, P. R., "Optically Controlled Lateral PIN Diodes and Microwave Control Circuits", RCA Review, Vol. 47, p.p.443-456, Dec., 1986.
- [7] de Salles, A., "Optical Control of GaAs MESFETs", IEEE MTT, Vol. MTT-31, p.p. 812-820, 1983.
- [8] Paoella, A., Herczfled, P. R., " Optical Gain Control of a GaAs MMIC Distributed Amplifier", *Microwave and Optical Technology Letters*, Vol. 1, No. 1, March, 1988, p.p. 13-16
- [9] Paoella, A., Herczfled, P. R., " Optical Gain Control of a GaAs MMIC Transmit-Receive Module" 1988 IEEE MTT-S International Microwave Symposium , New York, N.Y.,May 24-27, 1988.
- [10]Selin, J. R., "Continuously Variable L-Band Monolithic GaAs Phase Shifter", *Microwave Journal*, Sept. 1987.
- [11] W. D. Jemison, T. Berceli, A. Paoella, P. R. Herczfled, D. Kasemset and A. W. Jacomb-Hood,"Optical Control of a Digital Phase Shifter", 1990 IEEE MTT-S International Microwave Symposium, Dallas Texas, May 1990.
- [12] Herczfled, P. R., Daryoush, A. S.,Kieli, M., Siegel, S., Soeref, R., "Wide-Band True Time Delay Phase Shifter Devices", 1987 IEEE MTT-S International Microwave Symposium, Las Vegas, Na., 1987.
- [13] A. Paoella, D. Sturzebecher and A. Madjar, P. R. Herczfled, "Optically Controlled GaAs MMIC SWITCH using a MESFET as an Optical Detector", 1990 IEEE MTT-S International Microwave Symposium, Dallas Texas, May 1990.
- [14] Daryoush, A. S. et al., "Fiber Optic Memory Loop", MILCOM 88, San Diego, Ca., Oct., 1988.
- [15] Paoella, A., Higgins, T., Herczfled, P., "An Optically Controlled Millimeter Wave Dielectric Resonator Oscillator", IEEE Seventh Annual Ben Franklin Symposium on Advances in Antenna and Microwave Tech., Philadelphia, Pa., March 13, 1989.

Directed Energy Warfare: Advances in Protective Armor Materials
Dr. Janet Perkins, Ms. Pearl W. Yip and Mr. Charles G. Pergantis,
U.S. Army Materials Technology Laboratory
Watertown, MA 02172-0001
Dr. John W. Walkinshaw
University of Lowell
Lowell, MA 01854-3698
Dr. Ralph Brewer and Mr. John M. Gale,
OptiMetrics, Inc.
Las Cruces, NM 88005-3253

INTRODUCTION

Since the discovery, development, and emergence of high-energy directed beams - particularly lasers - which can almost instantaneously deposit this energy on a structure at some distance from the laser-producing device, the United States must explore both the potential use, and protection against unfriendly use, of this new device. It is powerful, readily directed, and could be a major threat in the event of military confrontation. Consequently, the military services, sharing their research results, have been exploring the survivability and protection response of the most advanced new materials to the effects of this new potential weapon.

A high energy laser (HEL) beam is a unique form of energy -- an intense, directed beam of electromagnetic radiation that may be invisible until it contacts solid material, in contrast to ballistic weaponry that carries kinetic and potential energy at a relatively slow rate. The investigation of the laser's effects was desirable and proved to be fascinating as well as productive.

This report will cover three topics:

- * a discussion of the fundamental effects of laser beams on materials;
- * our experimental study to determine the optimal selection among a class of protective materials;
- * factors involved in physical behavior and production of such a laser-hardened barrier material.

A laser produces a directed, coherent, narrow, beam of electromagnetic energy capable of transporting enormous energy a considerable distance with the speed of light; its fundamental action on a barrier material should be examined. During interaction with a material, the beam can be reflected, scattered, transmitted, and absorbed, producing photophysical and photochemical effects. The first three types of interaction produce minor damage; the last alters and can destroy a structural material. Hence our intent in seeking a protective material has primarily been to investigate the last effect.

RADIATION/SOLID INTERACTION

It has been known that large blocks of graphite can absorb, store, and slowly convert and release large amounts of radiant energy. Carbon, the element with the highest melting point, has extraordinarily strong interatomic and resonance-delocalized bonds, hence its use in nuclear reactors and jet engine throats. With a thin coating of the refractory metal tungsten over a 1/4-inch thick piece of isotropic graphite, we had observed a significant increase in resistance to burn-through by a medium power laser. This material was a variant of the very resistant but much more costly TBR, a "tungsten bearing resin" coated carbon cloth laminate, carbonized and graphitized under pressure, which at the time of the experiment was the "hardest" material known to minimize laser damage.

When a powerful laser beam interacts with a thin metal layer deposited on a carbon block, the metal provides a reflective surface to divert the beam. Only a small fraction of the beam is transmitted (about 5 percent), and this is predominantly absorbed by the graphite substrate where its action activates the bonding electrons from lower to higher molecular orbitals in femto (10^{-15}) seconds. The principal immediate effect, with continuous radiation, is to confine the electrons in more localized and, if the beam is powerful enough, into antibonding orbitals releasing carbon atoms (C_1). The remaining absorbed energy is then dissipated by both radiative and radiationless mechanisms to produce light, heat, and chemical change, the effects diagrammed in Figure 1. Even as this initial activation continues, relaxation sets in, but at a slower millisecond (10^{-3}) to second rate, converting a fraction of the potential energy to kinetic energy in the bulk carbon and the new C_1 atoms. The ultimate effect varies with the intensity and duration of radiation from no apparent effect, except for some degree of heating, to light emission, to phase change, and to weight loss or ablation.

Tungsten (W), being the highest melting metal and only the second highest melting element, was a reasonable choice for these coatings. However, as laser action continues with multiple photon absorption, the

carbon atoms, with weakened self-bondings, begin to diffuse into and through the body-centered cubic structure of the tungsten coating as well as laterally and ahead of the beam through that part of the carbon under the spot irradiated. The net effect on the metal is to weaken the W-W bonds and lower the melting point of tungsten some 600°C . This modified coating just under the laser beam, however, has strong W-C bonds and forms a stable liquid skin that reflects, reradiates, and is permeable to more and more released migrating carbon atoms. If quenched by suddenly shutting off the laser beam, it solidifies to a mixture of one or more tungsten carbides (W_2C , WC^{1-x} , and WC) and possibly agglomerated carbon atoms. Figure 2 represents a detailed phase diagram of a tungsten-carbon system.

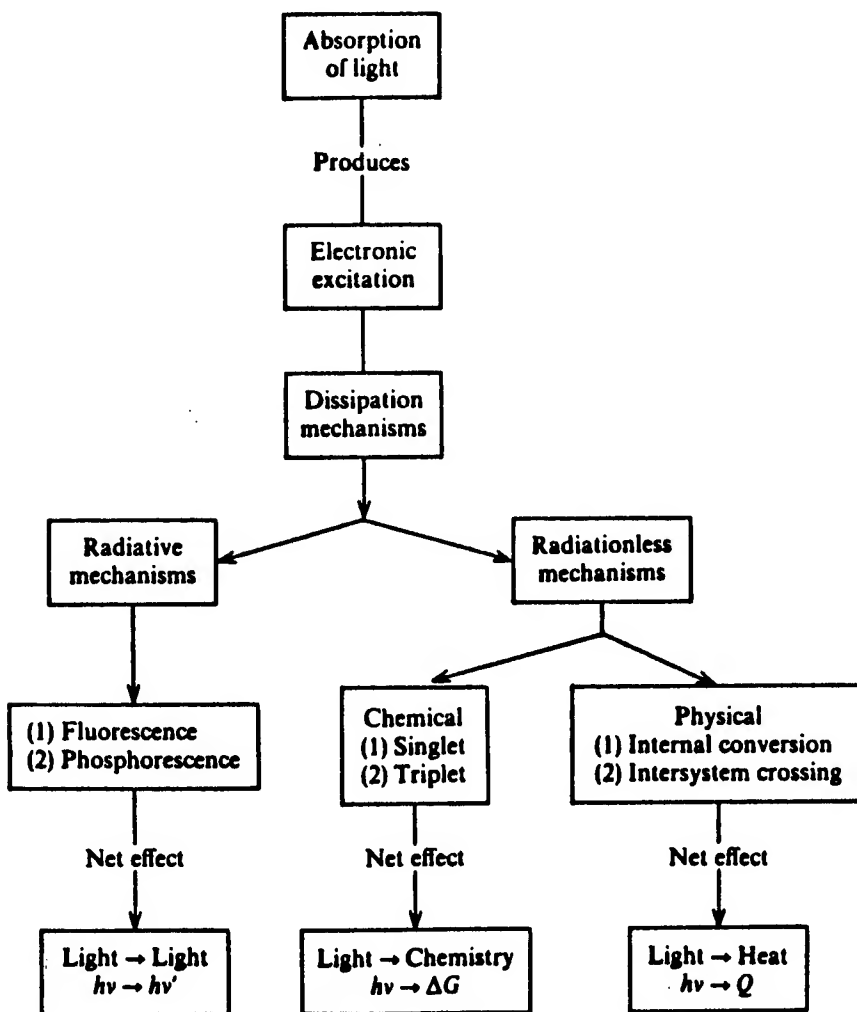


FIGURE 1. Flow chart showing photochemical and photophysical effects of irradiation on materials.

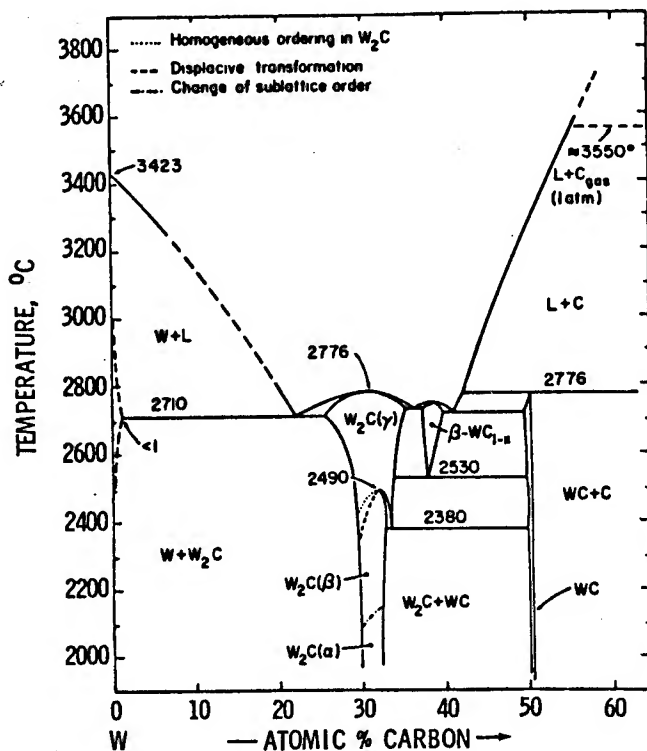


FIGURE 2. Phase diagram of a tungsten carbon system.

The distribution of the three carbide species depends upon the degree of diffusion, the stability of the solid forms, and the temperature reached. The defect form WC_{1-x} is stable and formed only above 2500°C, W_2C , only above 1300°C: WC has a peritectic melting point at 2776°C and is stable below. Crystallographically, the solidified W_2C and WC_{1-x} have been modeled with octahedra of W atoms, half or more of which enclose C atoms, which migrate from one tetrahedral vacancy to another^{1,2}.

The absorbed radiation first vaporizes the "softer" carbon in the underlying isotropic graphite producing an increasingly hot vapor which expands the tungsten and carbon skin into a stable bubble through which the C atoms continuously diffuse but which shields the underlying structural carbon.

As the carbon underneath the film vaporizes, creating a cavity as well as increasing pressure on the molten film, the latter eventually bursts, and if very cohesive, forms a rim of liquid and droplets on the upper wall of the cavity, exposing the underlying carbon to more direct laser action. Frequently, as with tungsten, the molten liquid wets the walls of the cavity, and runs down slowing the escape of carbon atoms and displaced particles of multiatomic clusters (Figure 3a). At this stage an ablation plume is frequently produced consisting of atoms, electrons, molecular carbon species, and loosened carbon clusters, much as occurs early on irradiation of the uncoated carbon substrate.

Shipkov et al.³ have described the accelerated vaporization of carbon in contact with the eutectic compositions metal-metal carbide and metal carbide-carbon, as well as recrystallization of graphite from a saturated solution of a refractory metal carbide. Vaporization slows down if a solid, higher melting carbide forms at the carbon eutectic interface; graphite recrystallization occurs on the surface of the

super-saturated carbide solutions obtained with very intense radiation. We have observed the latter when the related TBR was irradiated with a high energy laser beam (Figure 3b) on short exposure, and again on



FIGURE 3a. Flow of tungsten carbide melt into cavity. Note removal of "softer" carbon below burn tip.



FIGURE 3b. Recrystallized graphite produced at high laser power.

prolonged irradiation at a lower intensity of a metal-coated carbon in the X-ray diffraction pattern of the carbide residue near the edge of a burn. A slender diffraction peak appeared adjacent to the broad 002 peak of carbon exactly at the 2-theta value corresponding to 3.35Å, the interlayer distance in a pure graphite crystal.

Our observations of the striking effect of tungsten-coated graphite led us to explore the behavior of other refractory metals, not only to look for a possible better coating for graphite, but to increase our understanding of the processes underlying the observed behavior which could impact on manufacturing cost. The ten metals that are refractory - titanium (Ti), zirconium (Zr), hafnium (Hf), vanadium (V), niobium (Nb), tantalum (Ta), chromium (Cr), molybdenum (Mo), tungsten (W), and rhenium (Re) - are all high melting, stable, and except for Re, form carbides known for their stability and mechanical hardness. All share the similar atomic structure of the d-shell next to the valence shell being less than or up to half filled with electrons. Figure 4 compares their phase diagrams at the solidus level.

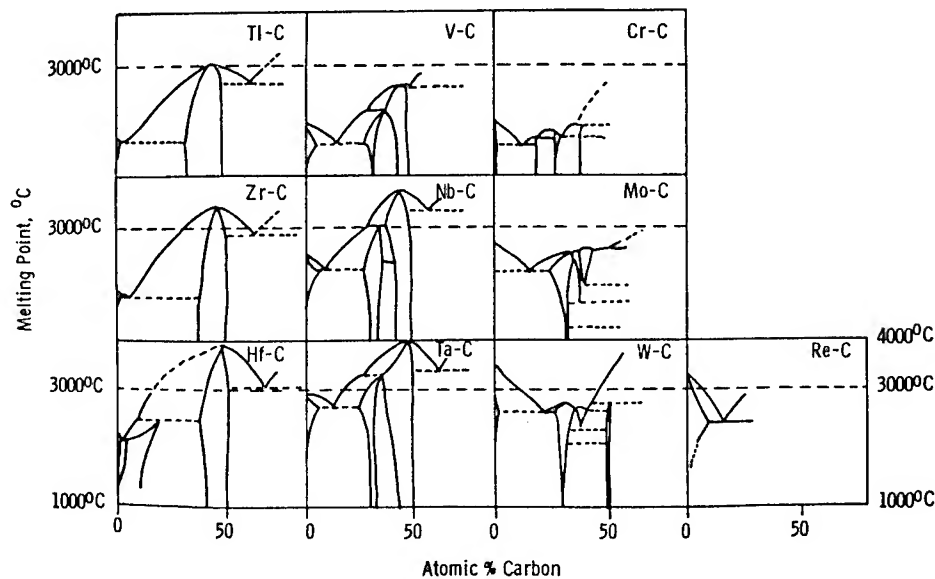


FIGURE 4. Phase diagrams of refractory metal carbon systems.

EXPERIMENTAL

A total of 61 samples and controls were irradiated in a large vacuum chamber with a high energy continuous wave CO_2 laser focused to provide a known flux on one square centimeter at the center of a sample either 2.5cm or 5.0cm in diameter.

a. Sample Fabrication and Characterization

Samples were fabricated on contract with a nominal 1-mil coating produced either by chemical vapor deposition (CVD), plasma spray (PS), or bonding free-standing film to the substrate. A few CVD and one PS sample were requisitioned with carbide, boride, or nitride coatings. The coated samples received were found to vary as much as 0.2 to 1.2 mils in film thickness, though not in a given set. Samples were machined from 0.65-cm thick graphite stock. Thermal matching was done by carefully examining the thermal expansion coefficients of the metals and various grades of isotropic graphite.

All samples were weighed on receipt and characterized by X-radiography for uniformity, X-ray diffraction for composition, and a microsectioning technique to measure coating thickness. The room temperature reflectivities of the samples were measured at $10.6\mu\text{m}$ with a CO_2 laser, courtesy of Dr. Michael J. Berry at HARC.

b. Apparatus and Diagnostics During Irradiation

Figure 5 diagrams the key features of the experimental apparatus used. The CO_2 laser beam, top-hat and continuous wave in character, was monitored by both a water-cooled calorimeter (A) and a much faster mercury-cadmium-telluride detector (B) before entering the vacuum chamber in which it was focused to irradiate one square centimeter at the center of the sample also mounted in the chamber. Each circular sample was held at four points by silicon nitride contacts in a rigidly positioned but demountable stainless steel frame. Each sample had side and back spring-loaded chromel-alumel (type K) thermocouples to indicate the heat flow in the carbon substrate. Front- and back-surface temperatures were monitored over the entire sample by two infrared scanning radiometers and the true front-face temperature was additionally followed by a spectrometer with a circularly variable filter focused on a small portion of the area being irradiated. Data from this measurement could subsequently be compared with Planck black body curves to give high temperatures without requiring a knowledge of emissivity changes. A rapid-scanning optical multichannel analyzer (OMA), spanning the near-ultraviolet through the near-infrared regions of the spectrum (240-1260 nm), was focused immediately in front of the surface of the sample to determine volatile species. Through windows in the vacuum chamber, a filter-protected high-speed camera and a color-recording video camera were focused on the front surface, and a similar video camera on the back surface to record events. The laser was set for predetermined times but could be aborted by the individual visually monitoring each run.

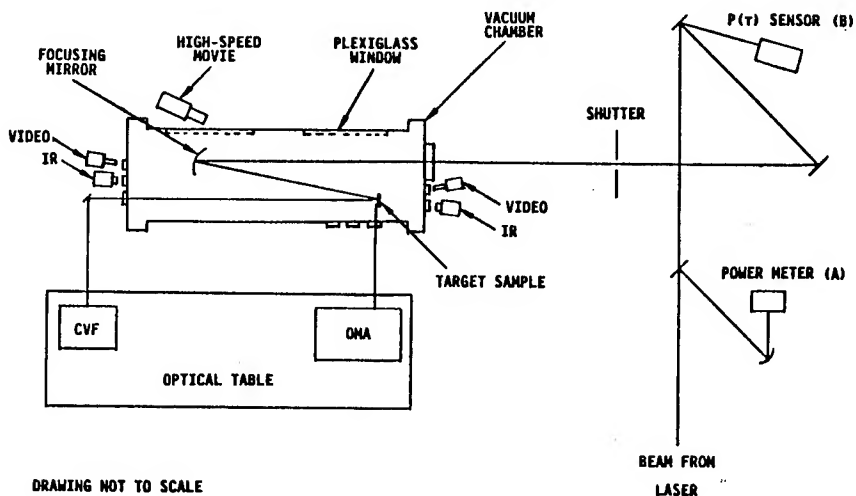


FIGURE 5. Experimental set-up.

The vacuum chamber was evacuated to 10^{-4} torr except between sample changes, when it was filled with argon to minimize air entry prior to reevacuation and avoid the presence of any air during irradiation. At the time samples were changed, and particularly after control runs, any accumulated soot was removed from the focusing mirror.

c. Post Irradiation Characterization

Immediately after removal the front and back surfaces of each sample were photographed and the samples were reweighed. Each sample was characterized again by X-ray diffraction both across the burn and near the edge to indicate the degree of carbonization in the two areas and any impurities displaced by the beam. Scanning electron micrographs (SEM) of selected samples, semi-quantitative energy dispersive spectral (EDS) analyses of microscopic features, and elemental mapping of areas were employed to examine more closely what had occurred during irradiation.

RESULTS

a. Weight Loss

The samples were listed in descending order on the basis of the total power received (before any burn-through) divided by the weight loss. This showed some of the materials to be clearly superior at the beam intensity studied.

b. Carbide and C₁ Formation

All but one of the irradiated deposits, metals or metal derivatives, were converted to carbides; the exception, Re, as noted above, is non-carbide-forming. As was seen in Figure (4), carbon lowers the melting point of Mo, Re, and W to about 3000°K, but raises that of the group IV and V refractory metals. Furthermore, these two groups of carbides and Re remain liquid over several thousand degrees and hence have high stability. The heat of vaporization of carbon to C₁ is very high (170 kcal/g-atom); hence any mechanism involving C₁ formation and escape is very endothermic, forming a large energy sink.

c. Late Exotherm

Two unexpected thermal effects were observed. At the end of a long irradiation of a molybdenum sample, a large exotherm occurred as the film broke, releasing an ablation plume. One such temperature profile is shown in Figure 6; it seems unnaturally high. However, Kirillian⁴ has

reported measured vapor pressures of carbon produced by stationary laser heating at 5000-7000°K, and found average heats of evaporation equal to 90 ± 6 kcal/g-atom. This supports our observation of a major exotherm in a front-face temperature profile for the latter part of the irradiation of a Mo-coated sample. The other unexpected effect is described in Section e.

d. Thermal Distribution

The true front face temperature and the corresponding thermogram of the back face of another refractory metal coated sample after 30 seconds irradiation without burn-through is shown in Figure 7. The center back of the sample has reached a temperature of 528°C. (The temperature at the edge of the back face is 350°C.) In contrast, the apparent front face temperature under the beam measures 5030°C. At this stage of irradiation, the sample had lost less than fifteen percent of its original weight within the burn area. The combination of these observations gives some idea of the thermal protection afforded by this system.

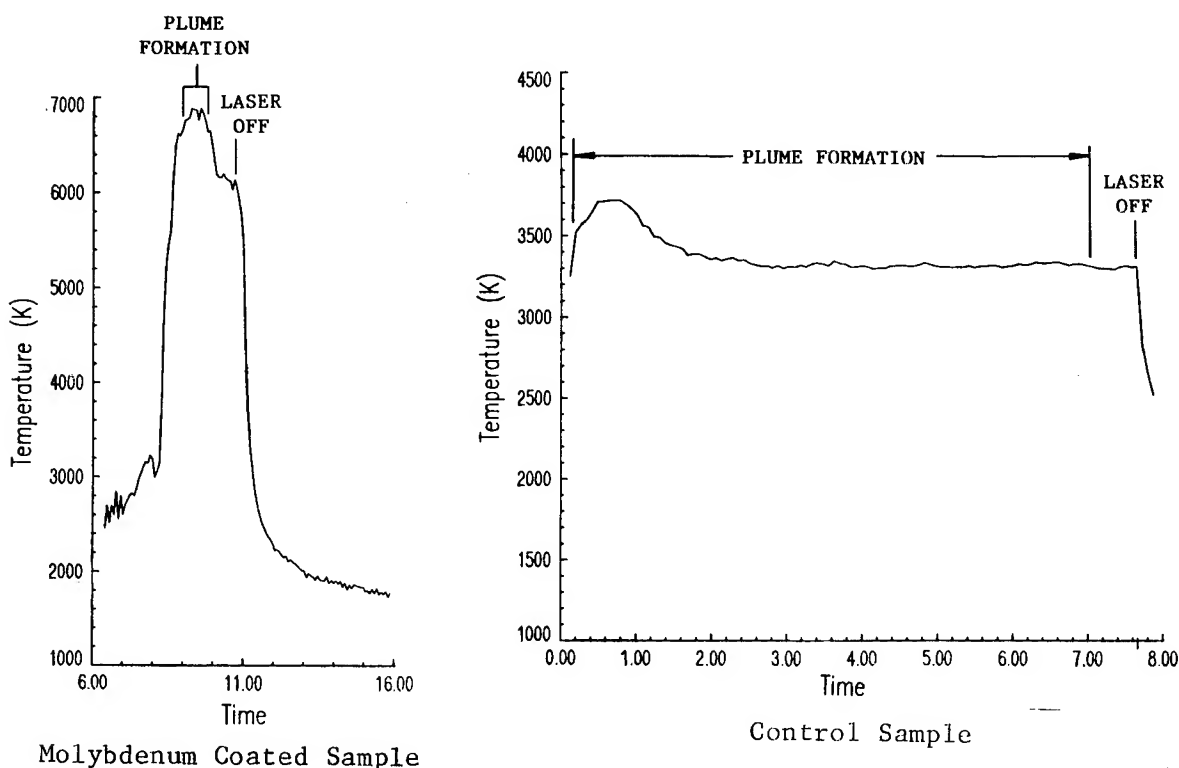
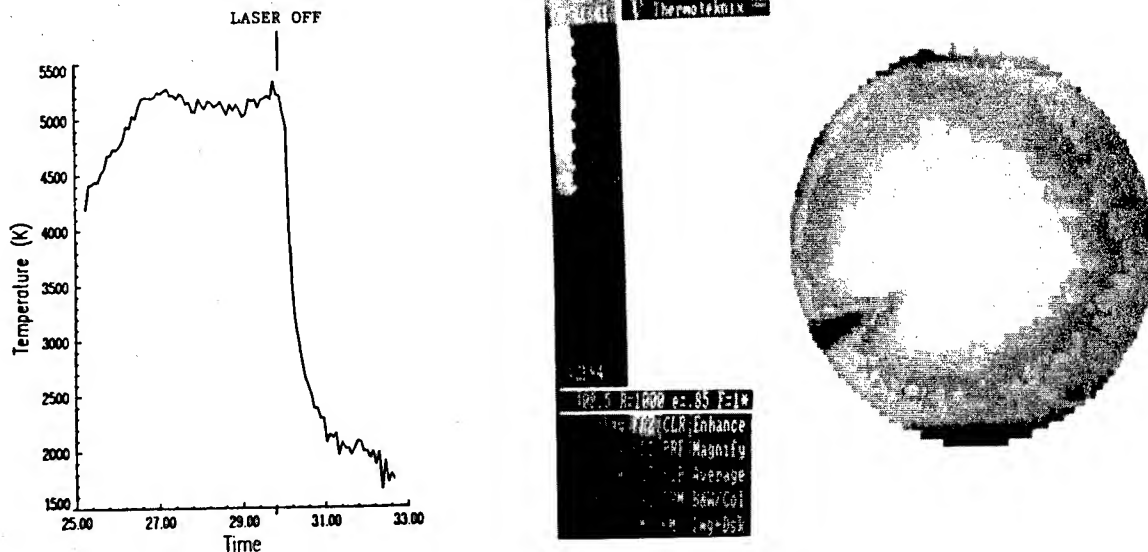


FIGURE 6. True front face temperature plots of a Molybdeum coated and a control sample.



a. Section of front face, temperature profile before burn-through of a refractory metal coated carbon.

b. Back face thermogram just before laser off point.

FIGURE 7. Superior sample of a refractory metal.

e. OMA Data on Plume Composition

An interesting observation was made on the temporal behavior of the C_2 content of the plume produced when the graphite controls were irradiated. The prominent C_2 Swan bands appeared, grew slightly, diminished, and then disappeared in about ten OMA scans as shown in Figure 8. This may be the result of (1) redissociation of C_2 , a secondary product, (2) increased production of C_1 , whose spectrum is out of range of our detector, and (3) delay in the formation of the ablation plume phenomenon, free-jet expansion described in reference 5.

Negligible C_2 is seen with coated samples until the metal/carbon film is breached.² Prior to this, any monatomic carbon, diffusing slowly through the metal/carbon layer, may dimerize on reaching the surface and be briefly visible until equilibrium is established.

f. Changes Evident from SEM's

The SEM's, Figures (9 a-f), provided visual evidence of a number of our conclusions from other techniques. Some of the films had melted during the runs, cascading downward into the cavities produced when the

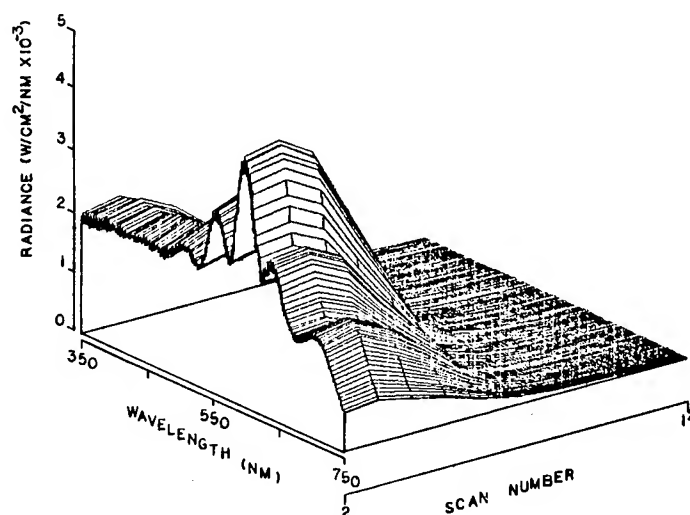


FIGURE 8. OMA scans of a typical control sample showing initial appearance then disappearance of Swan bands of C₂.

blistered skin broke (9a). Prior experience has shown that this light element can be boiled out initially from samples during laser irradiation. The Re film was clearly more molten (9b, 9c); the Ta film clearly stiffer (9d). The latter was thick and had rather large pores. The Nb film melted and flowed; although some studies have rated it highly as a barrier material, at this laser power the film seems too weak. The SEM of the sample shows escape holes by the rim for impurities we observed by EPS. A very interesting feature in both MoB and W samples

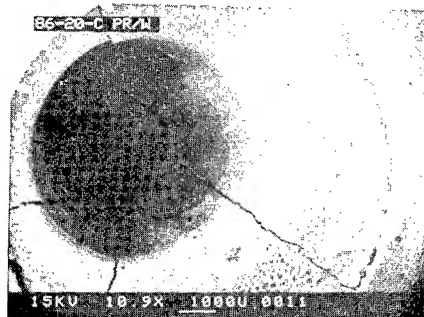


FIGURE 9a. Flow of tungsten (carbide) into cavity.

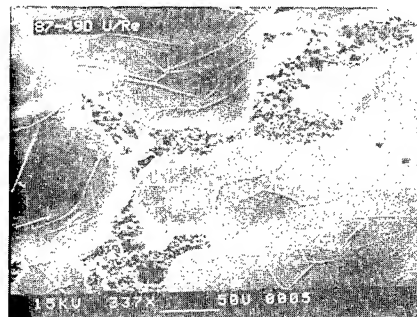


FIGURE 9b. Molten rhenium skin after irradiation; network is collapsed skin.

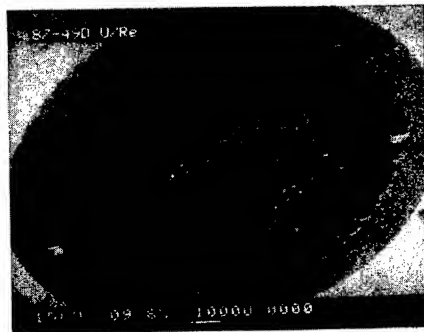


FIGURE 9c. Rhenium coated sample after burn. Note absence of cracks.

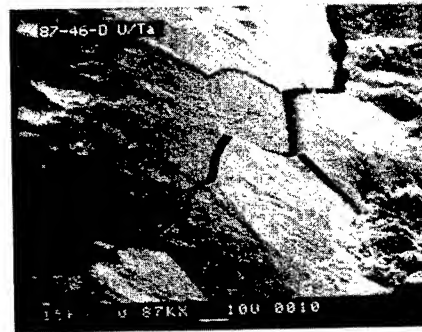


FIGURE 9d. More solid, brittle tantalum (carbide) sample after burn.

are tubular "pipes" pushed up from the carbon base on the side of cavities by escaping gas, and draped with the residue of the burst bubbles (9e). The "floor" of one Mo sample irradiated for an extra long time also shows clearly that the carbon substrate had become semi-solid (9f). Ta film tended to crack and crystallize in large grains, the probable cause of its poorer behavior at this power level. Other SEM's show retraction of carbide molten blisters, coalescence of cooled molten film pulling it away from unirradiated, unmelted coating beyond the "burn".

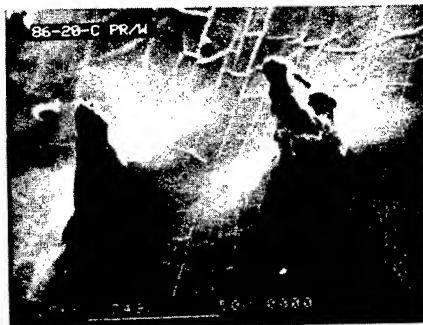


FIGURE 9e. Carbon "pipes" below collapsed tungsten (carbide) skin; grid lines are folds.

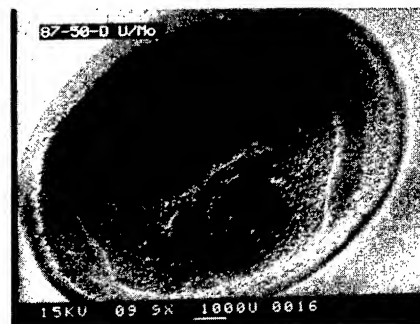


FIGURE 9f. Semisolid carbon substrate below split molybdenum (carbide) skin.

The lighter, lower melting refractories provided little protection at the power used; the heavier Hf and Ta films tend not to melt sufficiently and to be brittle.

CONCLUSIONS

The important factors for good barrier action appear to be the:

- (1) initial reflection of the laser beam by the solid or melted surface coat;
- (2) the strength and cohesiveness of this film over a wide temperature range;
- (3) the permeability of the metal (or metal carbide) film to neutral carbon atoms;
- (4) capability of the film to wet both amorphous and crystalline carbon, initially vaporizing the former, by transferring both radiant and subsequently-produced heat energy; and
- (5) optimal thickness of the film, which reflects, forms blisters which protect the subsurface, and transmits radiation, to photochemically activate the electronic structure of the solid. The latter subsequently produces reradiation, heat, melting, and vaporization as well as chemical changes. The atomization of the carbon structure is a tremendously endothermic process, diminishing the damage of that part of the laser beam not reflected or scattered.

A large graphite structure aided by a thin metal coating is light weight, capable of reflecting, conducting, storing, and reradiating energy from a powerful beam. It may serve as a protective and structurally resistant material under such adverse conditions as found on a modern or future battlefield. Yet, compared to reinforced TBR, it may have lower ballistic performance.

The prime advantage we see in coating a machined isotropic carbon part is the inherent labor and cost saving process of a spray or a CVD deposit over the successive cloth impregnation, cure, molding, post cure, and high-temperature and pressure heat-treatment of the alternate product. There appears to be little loss in laser hardening capacity.

Several areas requiring further investigation are the use of bimetallic films, as well as the use of lasers with different wavelengths.

REFERENCES

1. I.I. Kovenskii, "Diffusion in Body-Centered Cubic Metals", American Society of Metals, Menlo Park, Ohio, 1965, pp. 283-287.
2. G.V. Samsonov, "Diffusion Processes in Refractory Carbides, Physical Materials Technology of Carbides", Ch VIII, Fizicheskoya Materialovedenuye Karbedov, Kiev, Naukova Dumka, 1974. (English Tranl.), FSTC-HT-720-83, Foreign Science & Technology Center, Charlottesville, VA, 1983.
3. N.N. Shipkov, V.I. Kostikov, Ye I. Neprochin, and A.V. Demin, "Recrystallized Graphite", Metallurgiya, Moscow, 1979, (English Tranl.).
4. A.V. Kirillian, M.D. Kovalenko, M.A. Sheindlin, and V.S. Zhivopistsev, 1985, 23 (4), 699-706. (English Tranl.) High Temperature Institute, Academy of Sciences of the USSR, Plenum Publishing, 1986.
5. M.A. Covington, G.N. Liu, and K.A. Lincoln, 1977, AIAA J., 15, 1174.

Ranger and Special Forces Research:
Implications for Low Intensity Conflict Training and Operations (U)

Robert J. Pleban, Dr.
*Thomas J. Thompson, Dr.
Patrick J. Valentine, Mr.
U.S. Army Research Institute
Fort Benning, GA 31905-0686

Technological advances in night vision equipment, electronic sensors, weapon system performance, and communication systems have made it possible to conduct around-the-clock military operations. Whereas in the past soldiers could often rest and regroup during the night for the next day's activities, military units are now frequently required to perform in a sustained yet efficient manner for periods of time in excess of that normally considered possible for an individual (Angus & Heslegrave, 1985). The increased duration, pace, and intensity of such operations can be expected to severely tax the individual's endurance and recuperative power.

While the most frequent continuous operation scenarios were envisioned to be that of US/NATO forces battling successive echelons of Soviet/Warsaw Pact troops without let up, the focus of recent scenarios has gradually shifted. Increasing emphasis has been directed to low intensity conflict environments where irregular forces and procedures play leading rather than supporting roles; and where battlefields, rear areas, belligerents, and bystanders all are frequently ill defined (Collins, 1987).

The creation of Special Operations Forces (SOF) units which are capable of functioning in the environments described above requires, above all, the development of motivated, highly trained and disciplined soldiers. The necessary skills and qualities for successful combat performance cannot be created after emergencies occur, they must be developed and well honed beforehand. In the last few years there has been an increased interest in the development of appropriate training and selection procedures to enhance Special Operations Forces capabilities and the utilization of available training resources.

For the past five years the Army Research Institute (ARI) Fort Benning Field Unit's Light Forces Readiness team has been involved in both basic and applied research directed toward enhancing the capabilities of Light Infantry and Special Operations Forces by improving performance sustainability, training, and training selection procedures. These efforts have focused primarily on the identification of appropriate work-rest schedules for ranger students, the selection of small unit teams, and participation in the development and administration of a multi-phase screening program designed to select qualified Special Forces candidates.

This paper will summarize the research performed in these areas, point out the relevance of this research for low intensity conflict training and operations, and identify areas for further study.

Ranger Training

Although capable of operating across the combat spectrum (i. e., high/mid/low intensity), Rangers play a vital role in low intensity combat operations. Ranger specialties include quick strike and shock action over short periods, such as raids, ambushes, interdiction, and temporary seizure operations deep in hostile territory (Collins, 1987).

Initial training for Ranger qualification is intense. Over an 8-week period soldiers are exposed to a series of sustained operations conducted over varying terrains (mountains, swamps, and desert) and climatic conditions. The Ranger course is divided into four sequential phases: Phase 1-Benning, Phase 2-Mountain, 3-Swamp, and Phase 4-Desert. Within each operational scenario, little if any sleep is obtained. Between these operations only very brief sleep periods are allocated. Recent observations of Ranger training clearly demonstrated the deleterious effects of extensive sleep loss over extended periods of time. The majority of soldiers entering the latter two phases of the course appeared to be droning according to senior officer observers. That is, the students seemed to be just "going through the motions", incapable of either processing or acquiring any additional information.

The pervasiveness of the droning phenomenon and its negative impact on learning and safety subsequently led to a reassessment of current sleep schedule directives for the Ranger course which had been set at 2 hrs of sleep every 24 hrs or 4 hrs every 48 hrs. Two concerns were raised by senior Ranger staff members. First, no objective data existed which reliably indicated how much sleep Ranger students receive during the course. Second, was the sleep obtained by the students sufficient to minimize unnecessary risks inherent in some of the more hazardous parts of the course, e.g., live fire exercises. To address this concern, Pleban,

Valentine, Penetar, Redmond, and Belenky (in press) conducted a preliminary field investigation to provide data on the amounts of sleep obtained by Ranger students by day and training phase, the time of day sleep occurred, and the recuperative value of the sleep.

Summary of results. Fifteen U.S. Army soldiers were outfitted with wrist-worn, solid state activity monitors (see Redmond & Hegge, 1985) and had body weight and physical measurements taken while undergoing the 58-day, four-phase Ranger training course. Records from the activity monitors indicated that the average daily sleep obtained by the soldiers was 3.2 hrs. Average daily sleep by training phase was: Benning - 3.5 hrs; Mountain - 3.9 hrs; Swamp - 2.6 hrs; and Desert - 3.0 hrs. Average bodyweight at the start of the course was 178.2 lbs with a bodyfat composition of 14.7%. At the end of the course, the soldiers' average weight had dropped to 169.7 lbs. The change in weight was statistically significant, $t(9) = 3.89$, $p < .01$. Bodyfat, however, was essentially the same at 14.3%. This suggests that the weight loss may have been due to a reduction in lean body mass, indicative of the stresses placed upon the soldiers during training.

Mean ratings of both tiredness and sleepiness were generally high throughout the course. While the absolute levels of fatigue were high, soldiers reported feeling less tired, sleepy, anxious, and irritable, following their brief scheduled sleep periods, suggesting that their reduced sleep regimen had some recuperative value. Although the duration of sleep was clearly insufficient for complete recovery, Pleban et al. found that for certain phases of the course, sleep was significantly ($p < .05$) more restorative when obtained during the night hours (midnight-0559) than during the evening hours. Although no specific performance measures were collected as part of the research design, 33% of the soldiers could not complete the last three phases of the course. This percentage was slightly less than the typical attrition levels reported.

Implications. The results of this research along with specific recommendations on the management of work-rest cycles were briefed to the commander of the Ranger Training Brigade and the Commanding General of the U.S. Army Infantry School. These recommendations were used to alter selected aspects of the Ranger Training Program. Although the fluid nature of Ranger training limits the types of research that can be performed, Pleban et al. (in press) showed that one of the critical stressors of the program, i. e., sleep loss, can be studied reliably and the results utilized to enhance the quality of training received by the soldier.

Subsequent to the research conducted with the Ranger Training Program, the U.S. Army Training and Doctrine Command's (TRADOC) Combined

Arms Training Activity (CATA) completed a special study of high risk/high stress training programs. The observations and interviews that were a part of this study included an examination of the Ranger Training Program (TRADOC, 1989).

Special Forces Assessment and Selection

In 1986, ARI-Fort Benning was tasked by the John F. Kennedy Special Warfare Center (SWC) to assess current selection procedures used in screening candidates for entry into the Special Forces Qualification Course (SFQC) and to provide recommendations for improving the overall selection process. The general opinion was that the overall attrition rate in the SFQC was too high (40%-50%). There were also concerns that some of the candidates successfully completing the SFQC did not possess the psychological makeup or temperament to function successfully as members of an operational detachment or A-team.

An operational A-detachment consists of a twelve man unit (Career Management Field 18 is limited to male soldiers) proficient in five fundamental areas: light weapons, demolitions, field communications, combat intelligence, and paramedical support. While these units are capable of conducting raids, ambushes, and sabotage, the A-detachment is best employed as a "force multiplier" which develops, organizes, equips, trains, and directs indigenous military and paramilitary forces in unconventional warfare, and foreign internal defense. Although units can perform such functions as an adjunct to nuclear, chemical, and conventional operations across the combat spectrum, they are commonly employed in low intensity conflicts of the types described above (Collins, 1987).

For three years, ARI provided research support and technical assistance (i. e., general selection guidelines, assessor training) to SWC in the development of a screening program (Special Forces Assessment and Selection - SFAS) to assess key qualities previously determined to underlie success in the SFQC and general suitability for inclusion as an active member of an operational A-detachment.

The current (SFAS) program is three weeks in length and was implemented in 1988. The purpose of the program is to evaluate SF candidates across a series of exercises tapping their physical endurance and stamina, ability to work in small teams, and team-relevant personal traits (leadership, motivation, responsibility, and stability), along with obtaining actual performance measures (work samples) of candidates' proficiency on such requisite SFQC skills as land navigation. The objective of the research (Pleban, Thompson, & Allentoff, 1989) described below was to determine the potential utility of selected paper-and-pencil

measures of intelligence (Wonderlic Personnel Test-WPI) and personality (Jackson Personality Inventory-JPI) as well as relevant biographical information in predicting the successful completion of Phase I of the SFQC.

Summary of results. Intelligence, personality, and biographical measures were obtained from 293 soldiers attending Phase I of the SFQC at Fort Bragg, NC. The relationships between two of the selected predictor measures, WPT and JPI, and the Phase I performance criteria, were assessed through a correlation-multiple regression strategy. Predictive utility was assessed using a discriminant analysis procedure.

Subsequent analyses showed that the WPT correlated significantly ($r = .29$, $p < .0001$), with overall Phase I status (successful-unsuccessful). Significant correlations ($p < .05$), were also obtained between Phase I status and four of the 16 JPI scales: Energy Level ($r = .22$), Anxiety ($r = -.12$), Risk Taking ($r = .16$), and Infrequency ($r = -.16$). The results further suggested that SF candidates who have had prior specialized training emphasizing land navigation, map reading, and patrolling (e. g., Ranger, Reconnaissance, Jungle Warfare) are likely to be much better prepared to complete Phase I successfully than those candidates who have not had such training. Preliminary analysis of the profile of the successful SFQC candidate suggested that the individual is one who is above average in intelligence, persevering, calm, enterprising, and is generally proficient in basic military skills such as land navigation prior to entering Phase I.

In terms of prediction, the single best predictor of Phase I status was intelligence ($R = .29$, $F(1, 270) = 25.02$, $p < .0001$). The best predictive model of Phase I status consisted of three variables: the WPT, Energy Level, and Risk Taking ($R = .36$, $F(3, 268) = 13.25$, $p < .0001$). When employed with a typical class sample such as the present one, the model was able to increase the total percentage of correct classifications by 11.5%. If these findings are replicated, the practical significance of these gains in classification accuracy could be substantial considering the costs involved in SFQC training alone.

Implications. While the results described above must be regarded as preliminary, they suggest that qualities (i. e., intelligence, energy level, risk taking, anxiety) relevant to SFQC success can be reliably assessed by paper-and-pencil instruments and employed to provide additional information about the SF candidate. The utility of these paper-and-pencil measures and the field exercises employed in the SFAS program (designed to provide behavioral assessments of leadership, teamwork, motivation, responsibility, and stability) in predicting performance on an operational A-detachment remains an empirical issue at this time. Research addressing

this issue has been proposed, and some preliminary data collection has occurred. The data is currently being analyzed.

Enhancing the Effectiveness of Long Range Reconnaissance Teams

While Special Forces units are predominantly involved in unconventional warfare and foreign internal defense missions, they also perform other critical missions such as long range, or strategic reconnaissance. Long range reconnaissance teams (4-6 man teams) are frequently forced to spend long periods of time deep within enemy territory while remaining undetected. While certain aspects of the mission can be boring or tedious, there is, nevertheless, the constant element of stress involved as the team tries to carry out its mission undetected by indigenous personnel, with team members often in close physical proximity for extended periods of time. Clearly, the compatibility of the team members' interpersonal styles must be considered if they are to work effectively under these conditions.

Renewed interest in long range reconnaissance units resulted in ARI-Fort Benning's participation in the identification of specific skills and qualities, relevant to enhancing these units' effectiveness. This work also has had some relevance for the development of the Long Range Surveillance Units that are part of the U.S. Army's Light Infantry Divisions. The focus of ARI-Fort Benning's work was not so much the identification of technical skills, but rather determining ways of enhancing small group compatibility. Using a model developed by Shutz (1958), exploratory research (Pleban, Valentine & Thompson, 1987) was conducted to examine the relationship between interpersonal compatibility, cohesion, and productivity.

Summary of results. Twenty four soldiers (four 6-man teams) enrolled in the nine week Advanced Land Reconnaissance Course (ALRC) at Fort Bragg, NC were monitored over a 10 day field exercise. The field exercise was designed to approximate all phases of an actual reconnaissance mission. The major objective of the mission was to gather information on any movement of equipment or personnel through a designated area of observation. The secondary objective was to complete the mission without being detected while moving into a designated area, while operating in the area, and while withdrawing from the area toward a predetermined extraction point.

Following the completion of the exercise, the soldiers were debriefed and team performance was critiqued. Next, both the evaluators and soldiers filled out two questionnaires designed to assess team effectiveness and cohesion. In addition, the soldiers also filled out the Fundamental

Interpersonal Relations Orientation-Behavior (FIRO-B) questionnaire to determine individual orientation toward group compatibility. Rank order correlations were computed between specific measures of group compatibility, cohesion, and team effectiveness. Although the sample size was too small to make conclusive statements, the results were quite consistent. The more compatible the group the more cohesive/effective the group was rated. Of the twelve rank order correlations computed, all were positive. The average size of the correlations was .76. Due to the small sample size, however, only two correlations reached statistical significance ($p < .05$).

Implications. The stresses imposed by combat and the missions described earlier on small groups require that group members function as a cohesive unit if they are to successfully complete their mission. While effective strategies exist for molding cohesive units (see Henderson, 1985), they generally take time to implement. It may however, be possible to accelerate the cohesion building process by carefully selecting and matching individuals on the basis of certain interpersonal qualities and individual needs, and assigning those individuals who "match up" to the same unit. While this may not be feasible for conventional units, it may be practical and necessary for certain SOF units with specific missions in low intensity conflict environments. The results, briefly described above, suggest that the construct of compatibility, as assessed by the FIRO-B, may be an important mediator of both small team cohesion and performance. Moreover, the FIRO-B may offer a cost effective means of assembling small groups with the capacity for rapidly developing into cohesive and effective units.

Future Directions

The tremendous changes the U.S. Army currently faces in the geopolitical arena suggest that more rather than fewer Low Intensity Conflict operations can be anticipated. The increased use of conventional Light Infantry forces as well as Special Forces Mobile Training Teams to represent national interests can be assumed. At the same time the impact of constrained resourcing will be felt Armywide in training development and fielding. The base of experience gained from previous light forces small unit assessment and training research can be used to enhance the Army's training and maintenance of combat readiness for low intensity conflict operations. Opportunities for exploratory and applied research exist which can benefit from the knowledge gained from previous work with Special Forces and Ranger soldiers.

To maintain combat readiness the U.S. Army must explore training methodologies and instructional technology which require limited resources and which are oriented toward application at the unit level. Research activities must give more consideration to the anticipated operational environment of the customer, the Army, as well. Current experiences indicate that any military activities, whether advisory or operational, will require familiarity with continuous operations capabilities in low intensity conflict environments. In the past such operations have proven problematic. Observations and research with selected U.S. Army units and host nation military forces represent an opportunity to develop training interventions that would enhance the Army's appropriate combat readiness capabilities.

A related issue for investigation is the requirement to maintain combat readiness while managing and manipulating training risks and stressors which commonly increase with high fidelity training programs. An appropriate objective of related research is to enhance individual and small unit adaptation skills to perform successfully in continuous military operations. Two of the U.S. Army's field manuals, FM 26-2 Management of Stress in Army Operations (1986), and FM 22-9 Soldier Performance in Continuous Operations (1983), have been prepared to assist commanders who must train their soldiers for the rigors of combat and who must then monitor performance indicators to determine the level of fatigue and safety risk within their units. Observations made with U.S. Army personnel engaging in Mobile Training Team and advisory operations can provide insights and unique opportunities to field test interventions designed to enhance resistance to the onset of combat fatigue in counterinsurgency scenarios.

Criteria for personnel selection and training should be directed at identifying soldiers capable of better interactions with host nation military counterparts. A review of the current training procedures used in foreign area officer development programs can be used with newly acquired information to investigate appropriate selection criteria.

In summary, future research must address strategies for improving personnel selection and training for combat readiness in ways that recognize resource constraints. Research must also consider the environment in which the Army will be operating. The challenges that face the training development research community are to remember the needs of the customer, the U.S. Army, and to meet those needs quickly and efficiently with cost effective training products that clearly support readiness for combat in the low intensity conflict environment.

References

- Angus, R. G., & Heslegrave, R. J. (1985). Effects of sleep loss on sustained cognitive performance during a command and control simulation. Behavior Research Methods, Instruments, & Computers, 17, 55-67.
- Collins, J. M. (1987). United States and Soviet special operations. Washington, DC: U.S. Government Printing Office.
- Combined Arms Training Activity (1989). High risk/high stress training special study final report (Report No. 89-3171). Fort Leavenworth, KS: Author.
- Department of the Army. (1983). Soldier performance in continuous operations (FM 22-9). Washington, DC: Author.
- Department of the Army. (1986). Management of stress in army operations (FM 26-2) (rev. ed.). Washington, DC: Author.
- Henderson, W. D. (1985). Cohesion: The human element in combat. Washington, DC: National Defense University Press.
- Pleban, R. J., Thompson, T. J., & Allentoff, H. L. (1989). Preliminary assessment of selected predictors of special forces qualification course success (Research Report No. 1539). Alexandria, VA: U.S. Army Research Institute.
- Pleban, R. J., Valentine, P. J., & Thompson, T. J. (1987). Enhancing small group cohesion and effectiveness in long range reconnaissance teams (Research Note No. 87-30). Alexandria, VA: U.S. Army Research Institute.
- Pleban, R. J., Valentine, P. J., Penetar, D. M., Redmond, D. P., & Belenky, G. L. (in press). Characterization of sleep and body composition changes during ranger training. Military Psychology.
- Redmond, D. P., & Hegge, F. W. (1985). Observations on the design and specification of a wrist-worn human activity monitoring system. Behavior Research Methods, Instruments, & Computers, 17, 659-669.
- Shutz, W. C. (1958). FIRO: A three-dimensional theory of interpersonal behavior. New York: Rinehart.

Acutely Enhanced Burn Recovery by Post-Event Non-Ionic
Surfactant Treatment in a Rat Model (U)

*Kent M. Plowman, COL, MC
James C. McPherson, III, Dr.
Paul W. Paustian, Dr.
Randall R. Haase, CPT, MC
Royce R. Runner, Mr.
James C. McPherson, Jr, Dr.
Dwight D. Eisenhower Army Medical Center
Ft Gordon, GA 30905-5650

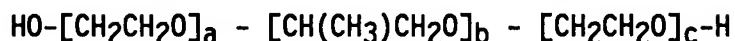
Most burn studies to date, which deal with the early stages of treatment, focus on fluid replacement. Early interventions aimed at rescuing marginal tissue have been of limited success or have required pre-burn administration. Inadequate basic knowledge of the pathophysiology of early burns has limited attempts at intervention in these crucial steps. The fundamental observations of vascular and perivascular changes of fibrin deposition, edema, leukocyte infiltration and erythrocyte extravasation are well recognized but poorly understood.

The current work examines a relatively new class of compounds in burn therapy. These are the Pluronic polyols, non-ionic surfactants, which have been shown to affect biological membranes and to have low toxicities. The Pluronic polyols are block co-polymers of ethylene and propylene oxides. Table I gives the relative mean molecular weights, biological half lives, and toxicities of the most commonly used members of this class of compounds.

Previous work in various laboratories have demonstrated beneficial effects at low concentrations of F-68 on growth of chick embryonic fibroblasts and hamster melanoma cells in tissue culture,¹ diminished chemotaxis and adherence of polymorphonuclear leukocytes in rats,² and improved rheological properties of erythrocytes.³ An externally applied gel formulation of F-127 has been previously used in pig burns as a protective barrier which can double as a chemotherapeutic reservoir.⁴⁻⁶ Nalbandian et al reported no differences at 24 hours in the histological analysis of the burned skin between experimental and control areas with a topical application of F-127.⁶ At 30 days the F-127 showed benefit only when gentamycin was present in the gel.

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

Table 1. Pluronic Polyols



Pluronic	Subunit Values			Mol. Wt. (Daltons)	Biological T½ (estimated hrs)	Mouse
	a	b	c			Acute IV LD ₅₀ Toxicity (g/Kg)
F-38	46	16	46	5,000	2	> 5.0
F-68	75	30	75	8,350	4	> 5.0
F-77	52	35	52	6,600	8	4.2
F-88	97	39	97	10,800	12	3.75
F-108	128	54	128	14,000	24	1.25
F-127	98	67	98	12,500	18	2.25

The present work built on our earlier studies in which Pluronics produced a tremendous endogenous elevation of the very low density lipoproteins (VLDL's) which are large lipoprotein molecules.⁷ Following burn injury, there is a marked increase in the blood vessel permeability resulting in an edema in which albumin molecules pass out into the edema fluid much more readily than do the larger gamma-globulin molecules. Since the VLDL's are also large protein molecules, they might be retained inside the damaged blood vessel like the gamma-globulin molecules and increase the colloidal osmotic pressure, thereby limiting the edema. In addition, other properties of non-ionic surfactants might play an important role in burns. Erythrocytes are known to exhibit damaged membranes after severe burn or after in vitro heating. We have shown that Pluronic F-68 protects heat damaged erythrocytes from hemolysis due to mechanical stress.⁸ Other potentially helpful effects of the Pluronic polyols in burns may be related to their ability to reverse the blood viscosity increase following burns.⁹

Methods: Pluronic F-127 was dissolved in distilled water to 12 mM concentration and adjusted to the isotonicity of rat blood (315 mOsm) by addition of crystalline NaCl. Each experiment used at least 10 adult male Sprague-Dawley rats (300-320 g) per group. The experimental group would receive F-127 at one or more times post-burn and the control group would receive equal doses of isotonic saline. The rats were anesthetized with Ketamine (50 mg/Kg) IM and Rompun (4 mg/Kg) IM, shaved to bare skin and placed in a standard template which exposed an area of the chest equal to 8% of the body surface area (BSA). The exposed skin area was dipped into a 70°C circulating water bath for 12 seconds to produce a third degree scald burn.¹⁰ The burn area was immediately outlined and measured for area. At 30 min post-burn, an intravenous injection was given of either F-127 or saline in a volume of 8 ml/Kg body weight via the tail vein. Thermography was recorded with an Inframetrics 520 unit

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

at 1.5 h post-burn (1 hr post-injection). The burns were photographed. Subsequent burn area measurements were performed at the indicated times under anesthesia. Erythrocyte deformability was determined with a Technicon Ektacytometer. The P₅₀ values were measured using a Hem-O-Scan analyzer. Blood samples were obtained from the tail vein. Animals were individually housed in clean stainless steel suspension cages with smooth wire mesh bottoms. No additional wound care was provided, i.e. no anti-biotic or surface ointment treatment was provided during the experimental period. Animals were observed for behavioral evidence of pain with provisions for analgesics if needed. None was required since the third degree burns were sharply demarcated and the burn areas were anesthetic. Some groups of animals were sacrificed 48 hours post burn. Other groups received repeated doses of saline (control group) or F-127 solution (experimental group) at the same dose as the 30 minute injection. These additional injections were given intraperitoneally. The repeated doses in both groups were given on day 4, 8, 12, 16, 20 and 24 (long term multi-dose) or on day 2 and 3 (short term multi-dose). Samples were obtained pre-burn, 6 hr, days 7, 14, 21 and 28. The animals in both the short and long term studies were ultimately sacrificed under anesthesia and autopsied with inclusion of a full-thickness skin biopsy which included normal tissue from the margins. Samples were fixed in 10% buffered formaldehyde for histological sectioning. Fixed specimens were sectioned and stained with hematoxylin and eosin by automated techniques. Blinded histologic comparisons were graded by at least two pathologists using a 0 to 3+ scale. Statistical analysis was performed using paired and unpaired Student's t-test on parametric data and a Chi square test on the histological scores.

Results: Surface areas immediately post-burn were not statistically different (24.8 ± 1.8 cm² for saline versus 26.0 ± 2.8 cm² for F-127 treated animals). A significant difference existed at 48 hours (16.6 ± 3.3 cm² for saline treatment versus 19.5 ± 3.4 cm² for F-127 animals, $p < 0.05$). This represents a 15% greater contraction of wound area at 48 hours in control animals. Gross appearances in wounds differed with greater edema (puffiness), hemorrhage and eschar evident in the controls over F-127 treated rats (Fig 1) at 48 hr post-burn. Table 2 lists the histological features scored in each group of singly injected animals at 48 hours. A marked shift toward less inflammatory responses is evident in the treated group (Fig 2). Gross autopsies were not different in visceral organs between groups. In single dose, short term multi-dose and long term multi-dose studies, wound areas were not significantly different from saline controls by day 7. However, the wound area remained consistently smaller, presumably due to the difference seen in wound contraction, during the first 48 hours through 28 days. In the single dose study, only 1 of 9 F-127 treated animals had an open wound remaining by 21 days whereas 5 of 10 saline controls exhibited open

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

Table 2. Histological Scoring Criteria

Vascular Necrosis (VN):	vessel walls showing a glossy homogeneous appearance, nuclear pyknosis-karyolysis with subsequent loss of cell structure;
Vascular Inflammation (VI):	vessel wall infiltration of polymorphonuclear leukocytes (PMNs);
Vascular Thrombosis (VT):	aggregation of blood factors primarily platelets and fibrin with entrapment of elements;
Fibrin Deposition (FD):	the deposit of fibrin in and around blood vessels giving the area a smudged, brightly eosinophilic, homogeneous appearance.
Perivascular Inflammation (PI):	infiltrate consisting mainly of PMNs and varying numbers of mononuclear cells;
Perivascular Fibrosis (PF):	fibrous exudate forming a thread-like, tangled eosinophilic meshwork around vessels with fibroblasts;
Edema (E):	excess fluid in the interstitial tissue;
Cellulitis (C):	inflammation of the cellular connective tissue;
Nuclear Fragmentation (NF):	karyorrhexis-leukocytoclasia; disintegration of leukocytes into "nuclear dust";
Erythrocyte Extravasation (EE):	the escape or discharge of blood from the vessel;

wounds indicating an enhanced wound healing in these animals. In long term multi-dose studies, there appeared to be a delay in complete wound closure where 7 of 9 F-127 treated animals had incomplete wound closure or scar at 28 days compared to 4 of 12 in the saline control group.

Histopathology of the burn wound area showed a lack of tissue necrotic indexes as outlined in Table 2 in the F-127 treated animals (Fig 3A). Vessel walls were intact, there was no vessel wall infiltration of polymorphonuclear leukocytes or surrounding fibrin deposition and no erythrocyte extravasation. There was some mild edema. In contrast (Fig 3B) the saline controls showed a marked increase in the same parameters with a predominance of perivascular inflammation, vessel wall

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

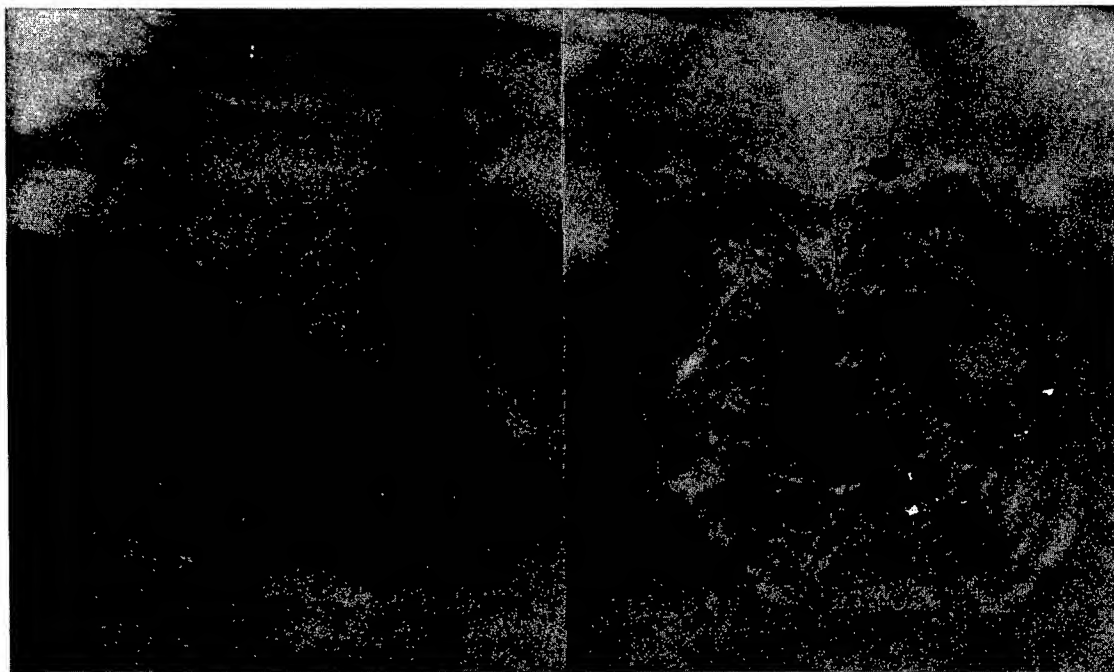


Figure 1. A third degree chest burn, 48 hrs after the burn in a single dose experiment, in a F-127 treated animal (left) and a saline control (right). Arrows denote area of eschar formation.

necrosis, vascular and thrombosis, fibrin deposition, cellulitis and edema. High power magnification of skin specimens from F-127 treated animals (Fig 4A) reveals intact vessel wall, individual erythrocytes within the vessel with no evidence of fibrin deposition, vascular thrombosis or vascular necrosis. In contrast (Fig 4B) saline treated controls consistently exhibited erythrocyte extravasation, vessel wall necrosis, vessel wall inflammation and vascular thrombosis. No fatalities occurred in either group.

Thermograms taken at 1.5 hours post-burn (Fig 5, 1 hour post-treatment) showed the burn area to be consistently 1°C warmer in controls (33°C) than in F-127 treated rats (32°C). This temperature difference is evident on both the thoracic and abdominal body surface area in each group. At 48 hours (Fig 6) the same 1°C difference existed but reversed with the saline animals averaging 32°C and the treated animals 33°C . The burn area is well demarcated and easily visible in

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

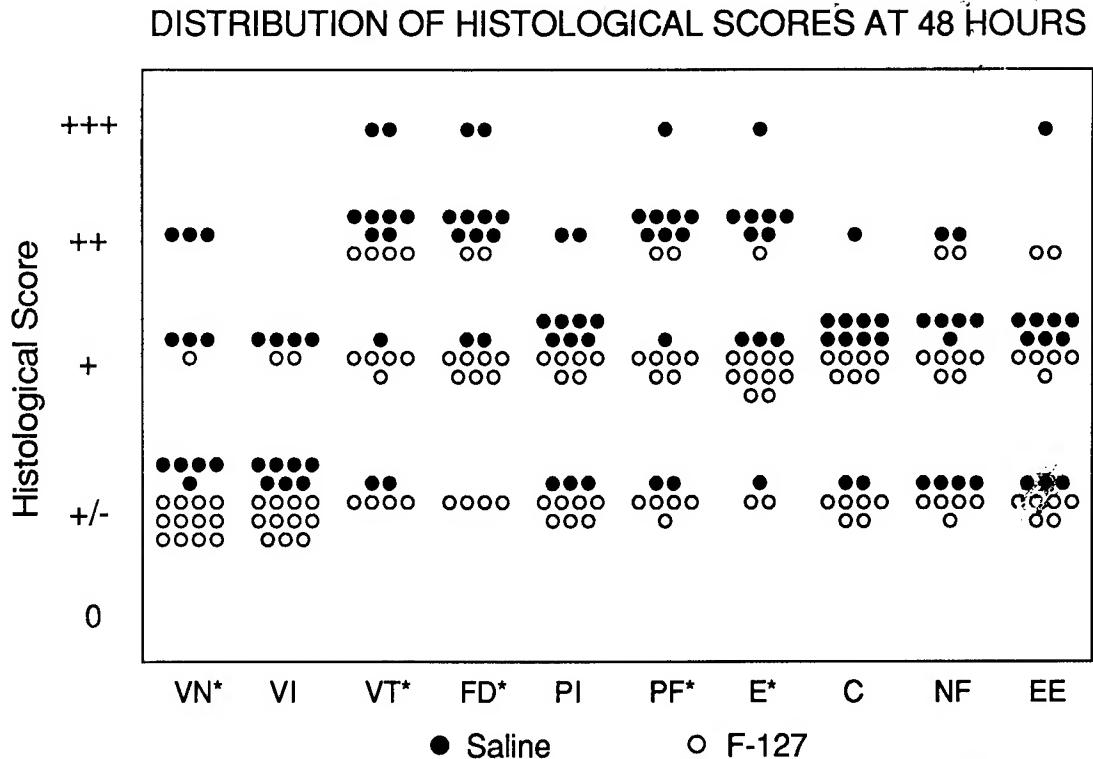


Figure 2. Distribution of histological scores at 48 hours. See Table 2 for abbreviations. Chi square, * $p < 0.05$.

the saline controls while the F-127 treated animals exhibit a uniform warmer body surface temperature of the thoracic and abdominal areas.

The erythrocyte deformability is not significantly affected in F-127 treated animals until about 6 hours at which time a highly significant decrease relative to zero time is noted (0.24 vs 0.21, $p < 0.001$) in the group sacrificed after 48 hours. This effect was not impressive in the long term multi-dose treated animals. In this latter group, however, an increase in deformability occurs by day 8 and persists through day 28 (0.26 vs 0.28, $p < 0.05$). A significant decrease in hematocrits compared to the saline controls is seen in the long term F-127 multi-dose study (45 vs 35 $p < 0.001$) on day 5 and the decrease remained through day 28. The values of P_{50} in rats receiving a 30% body surface area burn showed a small, but significant, shift to the right in F-127 treated animals at 1 hour post-treatment (34mm vs 39mm, $p < 0.01$).

Discussion: The significant observation of the present studies is that Pluronic F-127 treatment as late as 30 minutes after a third degree burn in a rat is capable of a significant reduction in several parameters of inflammatory damage to the tissues in the periphery of the burn area.

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

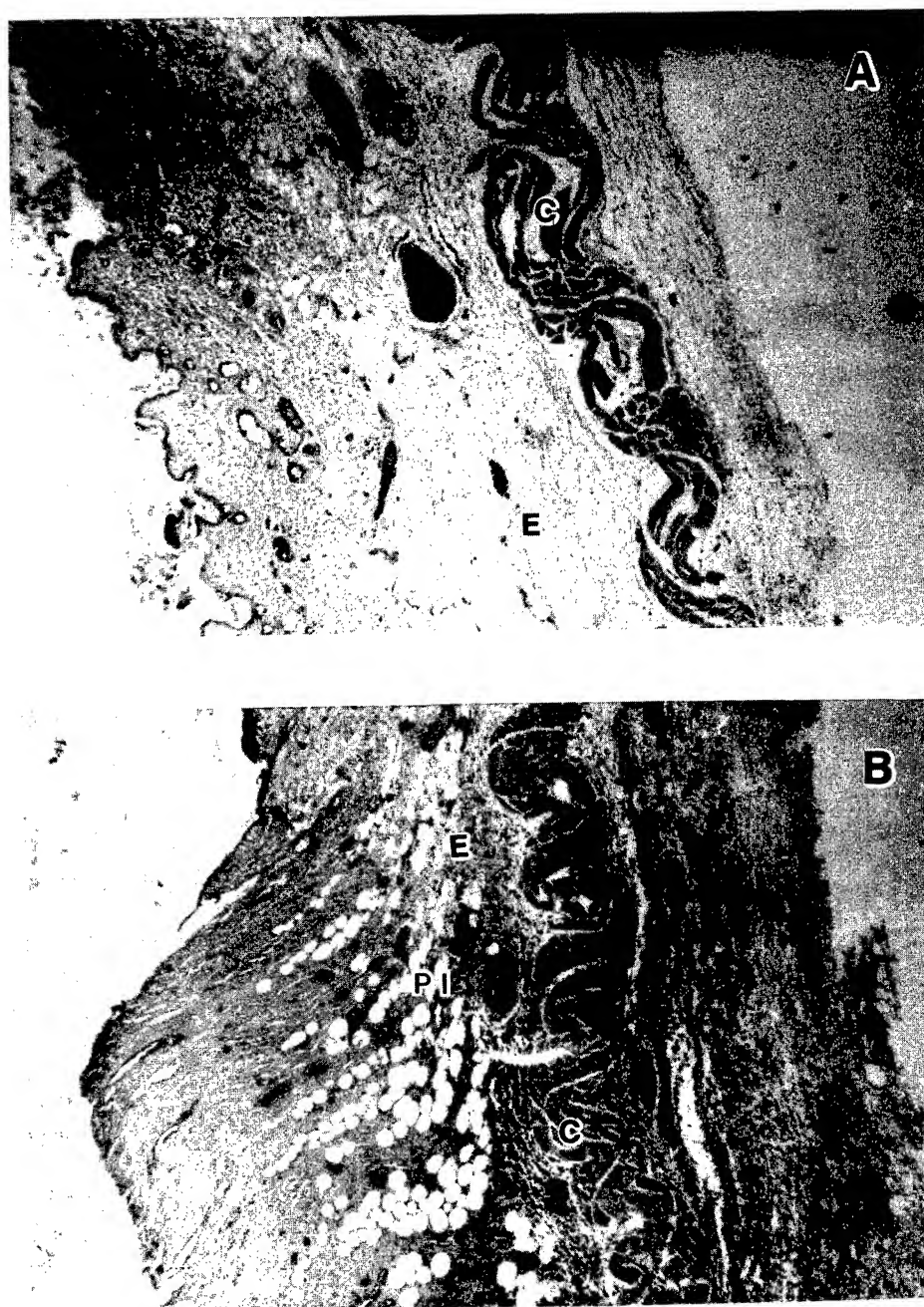


Figure 3. Photomicrographs of a cross sectional area of an excised third degree skin burn at 48 hours. F-127 treated animal (top, A) and saline control (bottom, B). See Table 2 for abbreviations. Magnification x 35.

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

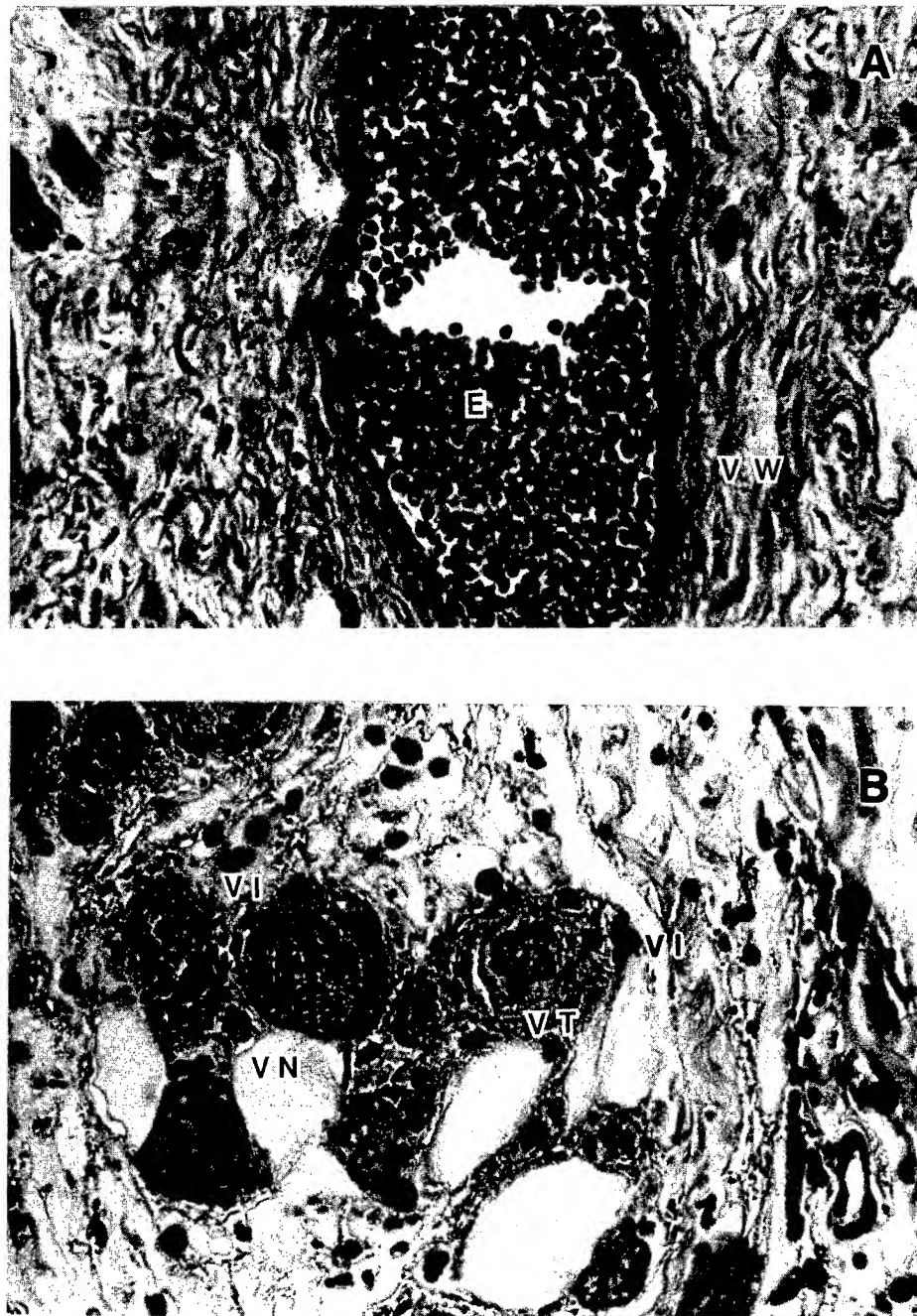


Figure 4. Photomicrographs of blood vessels in F-127 treated animals (top, A), in saline treated controls (bottom, B) at 48 hours post burn. See Table 2 for abbreviations. Magnification x 350.

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, MCPHERSON

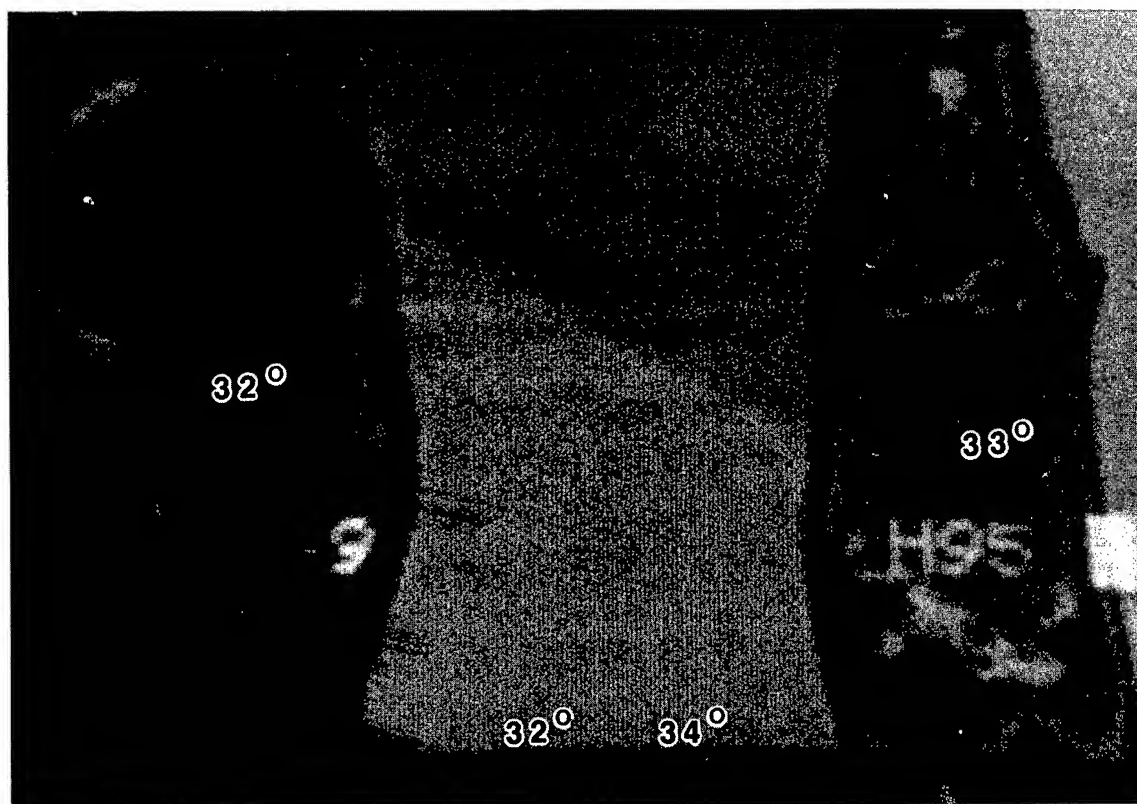


Figure 5. Thermograph taken at 90 minutes post burn (one hour post treatment) of a F-127 treated animal (left) and a saline control (right). Note the decreased temperature in the burn wound area of F-127 treated animal.

This is manifest in a variety of histological parameters evident 48 hours post-burn, especially in reduced edema and in perivascular and vascular damage. It allows for improved microvascular blood flow to the threatened underlying tissue structures as well as to the skin in the margin of the burn that is not irreversibly damaged.

As important as these observations are, the thermographic findings are especially intriguing in that they imply a mechanism which is exerting its effect within an hour. The one degree centigrade lower temperature in the wound area in treated animals is consistent with the hypothesis of improved microvascular blood flow allowing for a more rapid dissipation of heat relative to saline controls. It may also reflect a diminished immediate inflammatory response through some as yet unknown interruption of the inflammatory cascade. The Pluronic polyols are not generally regarded as anti-inflammatory drugs. The effect seen in these

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, MCPHERSON

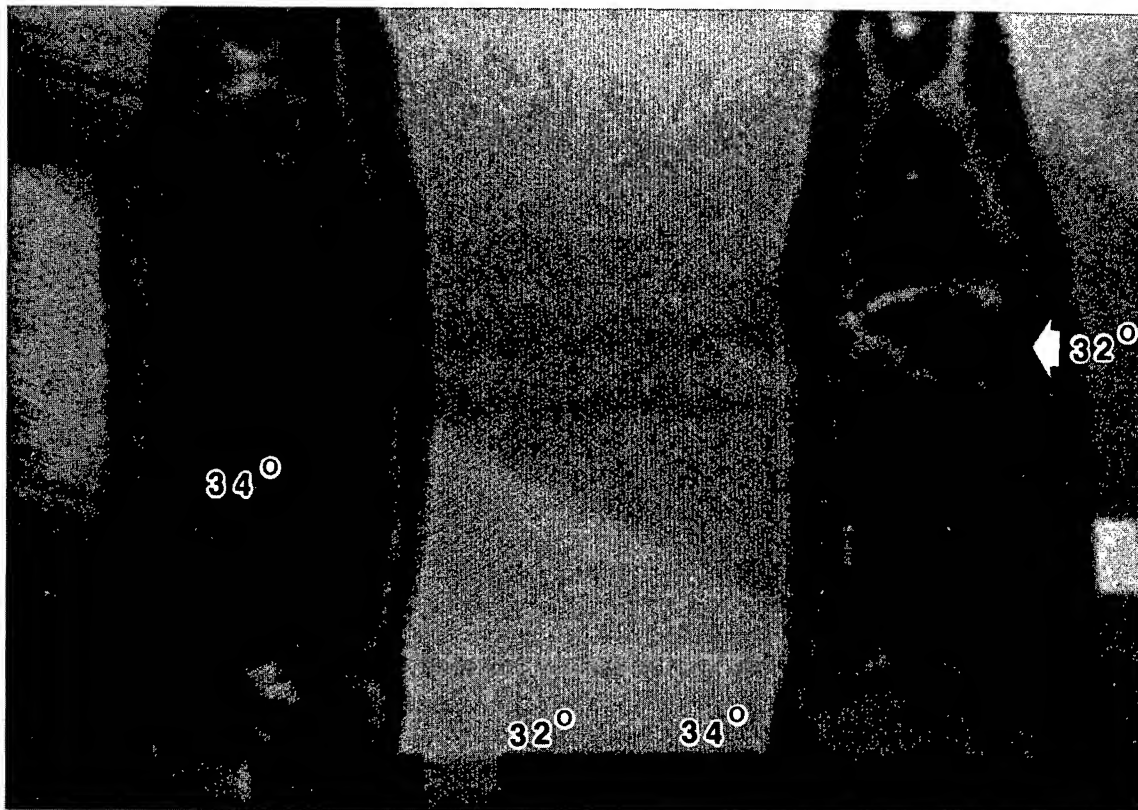


Figure 6. Thermograph taken at 48 hours post burn of an F-127 treated animal (left) and a saline control (right). Note the increased temperature in the burn wound area of the F-127 treated animal (arrows).

studies suggests that a post-injury inflammatory response can be interrupted as much as 30 minutes later. We do not know whether the 30 minutes is the optimal time for such an intervention since the time was chosen arbitrarily to reflect the interval needed between injury and practical intervention by paramedics. We had reasoned that an effect seen this late would have useful potential and have not examined the optimum time course. The present work would suggest that an earlier intervention with F-127 might have even more dramatic effects by further reducing the vascular damage.

The one degree warmer wound at 48 hours doubtlessly reflects the better microvascular flow to the wound area with less edema to obscure the radiation of heat. This more superficial blood flow allows for an orderly removal of necrotic tissue irreversibly damaged by the original

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

burn. These results would agree with Sowa¹¹ who used thermography to estimate burn depth through a correlation with clinical data based on dermal sensitivity, color variation and elasticity of burned skin. Stein, et al, ¹² correlated vascular proliferation changes during wound healing by means of thermography. The wound contraction data would again support the conclusion that more viable tissue remained in the treated animals in that less contraction was required to cover the wound. Contraction of the burn wound is a normal repair response through moving existing tissue at the wound edge to cover the burned area and does not represent the formation of new tissue.¹³

The blood rheology studies provide further evidence of beneficial effects from F-127. A small but significant in vitro effect on erythrocytes included immediate decreased deformability, decreased mechanical fragility, and a right shift in the P₅₀ which would bring about better oxygen delivery to the tissues. Erythrocyte deformability effects are not seen until six hours post-injection in animals given F-127 which suggests that those we observed are not due to an immediate effect of the F-127 on the erythrocyte membrane. The hematocrit between F-127 and saline control animals was significantly different on day 5 and thereafter, but was a factor only in the long term multi-dose study. The erythrocyte deformability effect was eventually reversed on day 8 of recovery in the long term multi-dose treated animals and remained increased through day 28. This increased erythrocyte deformability might reflect a tendency towards newer younger cells by that point in time because younger cells have increased deformability and removal of heat damaged erythrocytes would occur by this time. The erythrocyte effects are of interest in themselves because some change occurs within cell membranes in response to F-127 treatment (i.e., mechanical fragility) and other effects require an in vivo phenomenon which takes up to six hours to manifest. We have no way of knowing currently whether the mechanism is related to that which reduces the inflammatory response. Williams, et al² has reported that F-68 inhibits the in vitro adherence of polymorphonuclear leukocytes and decreased chemotaxis toward zymosan-activated serum in rats. F-68 also protected rat lungs from intratracheal bleomycin with reduced lung weights and fewer leukocytes. F-68 has also been shown to reduce pulmonary edema in experimental fat embolism syndrome. While the inflammatory response mechanism is suggestive, the change in deformability of the erythrocyte occurs only after a visible lipemia occurs.

Conclusions: The significance of this research is that the intervention can be at the site of the injury and provide some protection prior to evacuation to a hospital situation where more definitive care such as intensive shock therapy and skin grafting can be implemented. In peace time situations military personnel are exposed to the same causes of burn injury as are civilians. Moreover, they are more at risk for

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

fires in confined areas such as tanks, ships and aircraft. In times of conflict military personnel are much more likely to receive burns from exploding bombs, flame throwers, and steam lines. We have demonstrated a significant improvement in the histopathological events which occur following third degree burns by early intervention with a non-ionic surfactant, Pluronic F-127.

References:

1. Bentley PK, Gates RMC, Lowe KC, dePomerai DI, Walker JAL: In vitro cellular responses to a non-ionic surfactant, Pluronic F-68. *Biotechnology Letters*, 11(2):111-114, 1989.
2. Williams JH Jr, Chen M, Drew J, Panigan E, Hosseini S: Modulation of rat granulocyte traffic by a surface active agent in vitro and bleomycin injury. *Proc Soc Exp Biol Med*, 188(4):461-471, 1988.
3. McPherson JC III, Ward DF, Kirby SG, Yancey AL, McPherson JC Jr: Pluronic polyols and rat blood cell membrane fragility. *Fed Proc*, 44:1043, 1985.
4. Schmolka IR: Artificial skin I. Preparation and properties of Pluronic F-127 gels for treatment of burns. *J Biomed Mater Res*, 6:571-582, 1972.
5. Nalbandian RM, Henry RL, Wilkes HS: Artificial skin. II. Pluronic F-127 silver nitrate or silver lactate gel in the treatment of thermal burns. *J Biomed Mater Res*, 6:583-590, 1972.
6. Nalbandian RM, Henry RL, Balko KW, Adams DV, Neuman NR. Pluronic F-127 gel preparation as an artificial skin in the treatment of third degree burns in pigs. *J Biomed Mater Res*, 21:1135-1148, 1987.
7. McPherson JC Jr, McPherson JC III: Experimental hyperlipemic agents-Nontoxic alternative agents to Triton WR-1339. *Proc Soc Exp Biol Med*, 172:133, 1983.
8. Kirby SG, Ward DF, McPherson JC Jr, Prior RW, McPherson JC III: The protective effect of Pluronic F-68 against heat damage to the RBC membrane. *Georgia J Sci*, 44:38, 1986.
9. Davies JWL: Causes and effects of oligoemia. In: *Physiological Responses to Burning Injury*. Academic Press: New York, pp 106-140, 1982.

UNCLASSIFIED

PLOWMAN, MCPHERSON, PAUSTIAN, HAASE, RUNNER, & MCPHERSON

10. Paustian PW, Haase RR, McPherson JC Jr, Nguyen TH, Ward DF, McPherson JC III: Decreased inflammatory reaction (enhanced burn wound healing) in non-ionic surfactant treated rats. *Ann Emerg Med*, 18:480, 1989.
11. Sowa A. Thermographic estimation of burn depth. *Acta Thermographia*, 4:30-33, 1979.
12. Stein LE, Pijanowski GI, Johnson AL, MacCoy DM, Chato JC. A comparison of steady state and transient thermography techniques using a healing tendon model. *Vet Surg*, 17:90-96, 1988.
13. Peacock EG: Contraction. In: *Wound Repair*, Ed 3. Philadelphia, Saunders Co. 1984, p38-55.

UNCLASSIFIED

Role of Ascorbic Acid and Tin In Maillard
Browning of Wet-Pack Fruits

*William L. Porter, Dr., C.P. Dunne, Dr., D. Nattress, Mr.,
E.D. Black, Dr., J. MacNeill, Ms., D. F. Grant, Mr., J. J.
Pignatiello, Mr., and D. Rosano, Ms.

INTRODUCTION

The U.S. military is currently evaluating conversion from freeze-dehydrated fruit in the ration Meal, Ready-to-Eat (MRE) to fruit wet-packed in tri-laminated retort pouches and thermally processed for shelf stability. Wet-packed fruit has shown greater acceptance in field tests and encourages greater water intake. However, in the initial trial accelerated storage tests at 100°F (38°C), fruit wet-packed at Natick in retort pouches was found to fail much faster than the usual experience with similar fruit commercially produced in unlined tin cans. Brown color, baked and caramel odors, and mushy texture were the conspicuous quality losses. Further investigation led to the development of an improved processing procedure for preserving flavor and texture; this process was used by contractors for the peaches used in this investigation. Currently, very few vendors can process fresh fruit and pack it in retort pouches. Use of individually quick frozen fruit (IQF) and fruit repacked from cans increases the number of possible vendors and the time available for processing. There was a need for methods to evaluate these two alternative sources of fruit for the MRE.

At the 1988 Natick Science Symposium, we reported¹ that nonenzymatic browning occurs in pasteurized fruit that is wet-packed in retort pouches and stored at the high temperatures and long times mandated by military storage requirements. We reported that the fluorescence increases, off-color, cooked odors and mushy texture that result may be effectively inhibited by pretreatment with cation exchange resin to remove free amino acids. This finding implicated either phenolic browning or the Maillard reaction between reducing sugars and amino acids as possible causes of the observed browning.² In addition, studies of quality loss in peaches and pears that were packed in retort pouches in 1987^{1,3} revealed that prominent quality loss producers were: 1) excess air in the headspace, 2) lowering of pH (≤ 3.5), and 3) comminution of the fruit.

Approved for public
release, Distribution
unlimited.

Color and flavor losses like those we observed in wet-pack, retort pouch fruits stored at elevated temperatures to simulate field storage conditions might result from several causes: enzymatic and autoxidative browning involving phenols, thermal/acid caramelization of sugars, or Maillard browning from reactions of reducing sugars or oxidized ascorbic acid with amino acids or proteins.² Lipid oxidation, although theoretically possible, would produce increased fluorescence like that observed but would give yellow colors and rancid odors instead of the red-brown colors and baked odors we found.⁴ Most fruits have very low lipid levels and high water activity that minimize lipid oxidation.⁵ Phenolic browning quenches fluorescence, does not produce baked odors, and is strongly inhibited by normal heat processing and pretreatment with ascorbic acid like that used on the trial fruit. The fact that cation exchange resins, which sequester amino acids, strongly inhibit the browning largely precludes simple sugar caramelization as a quality loss factor for our pouch-packed fruits. Our search, therefore, concentrated on the Maillard reaction.

To monitor this reaction under various treatments, there was a need for rapid, sensitive, reliable, objective measurement methods, which would permit determination of the causes of quality loss and also allow prescription of preventive measures. We settled on front-face measurement of fluorescence (in contrast to the right angle viewing of solutions). Front-face measurement allows a rapid, sensitive and reproducible gauge of the Maillard reaction.^{1,2} Furthermore, using multiwavelength excitation and computer analysis, it promised to give the needed discrimination between sugar-amine and oxidized ascorbic acid-amine browning. Fluorescence has not been widely used in browning studies, because the fluorophore compounds are polymeric and often insoluble. They are usually colloiddally dispersed and therefore lost after clarification by filtration prior to studies of solution fluorescence. We have shown¹ that front-face sensing of fluorescence of the moderately turbid supernatants of homogenized, centrifuged fruit suspensions at an angle of 22° to the excitation beam gives good reproducibility, providing that one measures internal ratios of fluorescence emission intensity from two excitation wavelengths. Unlike the absolute emissions in front-face fluorescence, these change little with particulate concentration or sample position, both of which have unpredictable fluctuations in front-face work. Reference 1 gives details of the validation of this method against other Maillard measures. Also discussed are the strong correlations of fluorescence changes with color and degree of polymerization of the Maillard product.

In the present study, this method was used together with high performance liquid chromatographic (HPLC) measurement of sugars, 5-hydroxymethylfurfural (HMF) and furfural (F); colorimetric

measurement of ascorbic acid; reflectance; and pH. The fruit studied was peaches wet-packed in retort pouches in 1988 from either fresh, individually quick frozen (IQF), or repacked fruit, the latter having been initially packed in unlined tin cans. Conspicuous among the results after storage for six months at 100°F were the pronounced quality loss in the fresh fruit, which had a very high ascorbic acid level from processing, and the much reduced loss in the repacked fruit, which contained about 30 ppm tin ions and very much less ascorbic acid.

MATERIALS AND METHODS

Front-face Fluorescence Method of Assessing Maillard Reactions

Details of the method as applied to browned casein slurries and to wet-pack peaches and pears have been described.^{1,4} As used on fruit, measurements are made on a SPEX FLUOROLOG 2 Spectrofluorometer, using front-face excitation of solutions or slurries and sensing of emission at an angle of 22° to the exciting beam. The instrument is used in the DC mode, slits 2.0, 1.5, 1.5, 1.0, with 500 volts setting on the photomultiplier. A 390 nm filter is placed in the emission path. Excitations are at 360 nm and 395 nm and emission wavelengths observed, although variable, are typically at about 460 nm and 520 nm in fresh peaches, the latter shifting to below 500 nm as browning progresses. An excitation spectrum is read from an emission setting of 470 nm. (Typical spectra are shown in Figures 14, 15 and 16.)

Thus, at each reading five measurements are normally made: emission intensity at the max from both 360 nm (A) and 395 nm (B) excitation, excitation peak wavelength (C) and excitation intensity at 370 nm (D) and 395 nm (E) with emission set at 470 nm. From these values, two ratios are computed: B/A, emission ratio, and E/D, excitation ratio, both of which have excellent reproducibility. Also of moderate reproducibility and thus reported herein are the two emission intensities (A and B), and the excitation wavelength maximum (C) for 470 nm emission. Emission intensities are standardized with quinine sulphate (QS), 1 ppm in 0.1N sulphuric acid but are reported as μ amperes (μ A). Under these conditions, 1 ppm QS normally reads $1.25 \times 10^{-2} \mu$ A.

Preparation of Fruit Samples for Fluorescence Measurements

Twenty grams of stored fruit previously drained of excess syrup on an aluminum screen are added to 10 mL of the syrup and 10 mL of deionized water. This mixture is blended 2 min at high speed on a Waring blender and the slurry centrifuged 30 min in a clinical

centrifuge. The supernatant is used undiluted for front-face fluorescence measurements.

Calculation of Means and Variability of Results

For each stored sample, duplicate preparations are made and, after initial fluorescence readings, these are mixed, recentrifuged and reread, to give a total of four readings per sample. Means and standard deviations of the five above-mentioned parameters are calculated. Reported herein are emission ratio (B/A), emission intensity (B) and maximum emission wavelength from 395 nm excitation, and maximum excitation wavelength from 470 nm excitation (C). In the figures, standard deviations are reported only for B/A and B.

Processing and Packing of Peaches for Simulated Storage Tests*

Clingstone peach samples from three sources were packed and processed in U.S. Army MRE retort pouches: 1) fresh, 2) individually quick frozen (IQF) and 3) repacked from previous unlined tin can packaging. The first two types were packed and processed by a California contractor in August 1988 and the latter were done in-house in January 1989. Sliced fresh peaches, after lye peeling and washing, were held in ascorbic acid dips before pouch filling and syruing. The ascorbic acid dips were designed to yield a postprocess concentration of 500 ppm. IQF peaches, which were bleached and dipped in ascorbate, were used after thawing and draining to remove excess ascorbate. Results reported below, however, indicate that after packing and equilibration there were wide differences between types, the fresh fruit containing over 1100 and the other types between 200 and 250 ppm. After removal from the dip, the slices were packed in a sucrose syrup, which, after processing, was targeted to yield 18-22% Brix for the entire contents. The pH of the syrup was adjusted to 4.0 with a citric acid buffer. Pouches were filled, sealed and thermally processed at 190°F (88°C) for 3-4 minutes, which previous studies showed as optimal.^{1,3} Appropriate quality control tests of sample pouches were made for pH, Brix, enzyme (polyphenol oxidase) activity and approximate ascorbate level. Headspace was maintained at less than 10 cc.

*Simulated storage denotes controlled storage at two temperatures, 40°F (4°C) and 100°F.

Repacked peaches were prepared in-house from commercial canned peaches initially packed in "ultralite" sucrose syrup. The fruit had been in unlined tin cans for no more than 6 months prior to repacking. Adjustment of Brix, estimated ascorbic acid level, pH and thermal processing were carried out as for the fresh and IQF fruit.

Fresh and IQF fruit were equilibrated approximately 3 months at 40°F, until pH, Brix and sugar content had stabilized. Repacked fruit was held until it had equilibrated with respect to pH and temperature. Then the fruit was placed in controlled temperature storage at 40°F for control and 100°F as simulated high temperature field storage. Analyses were performed at this time, at 3 months and at 6 months.

Ascorbic Acid Analysis

A standard colorimetric method was used, based on 2,6-dichlorophenol-indophenol (DCIP) dye reduction.⁶ Standard ascorbic acid solutions are prepared in citric acid buffer and 3% metaphosphoric acid. Ethylene diamine tetraacetic acid (EDTA) is added to prevent interference from ferrous, cuprous, or stannous ions. To prepare a standard curve, 5 mL of standard is added to 5 mL of DCIP and absorbance is read at 520 nm, after 10 seconds, 20 seconds and after the addition of a grain of ascorbic acid. For the fruit sample, the entire pouch contents are blended 20 seconds in an Oster blender. A 5-10 gram sample is diluted in citric and metaphosphoric acid buffer, filtered and read as above. A computer program calculates ascorbic acid concentration in the sample homogenate. However Maillard sugar-amine browning-generated reductones, if present, may interfere in this method.

Sugar Analysis by HPLC

The peach homogenate from the ascorbic acid assay was used after centrifugation for 20 minutes at 10,000 rpm for pH, Brix and HPLC analysis of sugars, 5-hydroxymethylfurfural (HMF) and furfural (F). The sample for HPLC analysis was diluted tenfold with distilled water and filtered using a Millipore Millex-GV 0.22 µ filter. Analysis for fructose, glucose and sucrose was done on a Waters HPLC using a Waters Carbohydrate Analysis Column and differential refractometer detection. The solvent used was acetonitrile: water (80:20) and the flow rate was 2 mL/min. Output was handled by a Spectra Physics Integrator which reports concentrations using specific response factors calculated from external standards. Total sugar values are reported as the sum of concentrations of glucose, sucrose, and fructose determined by the integrator.

HMF and F Determinations by HPLC

A modified method ^{7,8} for HMF and F used dilute sulphuric acid (0.001N) as solvent at a flow rate of 0.5 mL/min., on a Varian Model 8500 HPLC with a Bio-Rad Fast Preservative Analysis column (30 x 4.6 mm). Detection was by a Kratos Spectroflow 783 UV/VIS Absorbance Detector set at 283 nm, outputting to a Waters 745/745B Integrator. Sample concentrations were calculated from a least squares standard curve of fit, of external standards at three different concentrations.

Reflectance

Reflectance measurements on peach homogenates (entire pouch contents) were done on a Hunter Color Difference Meter, Model D-25-9, in 50 x 8 mm Millipore covered plastic petri dishes, flush-filled with the fruit centrifugate after decanting supernatant. Appropriate white and grey plate standardization was carried out for each set of readings, which included Hunter L, a, and b.

RESULTS AND DISCUSSION

Retort Pouch Peach Storage Simulation Study

This part of the study compared quality loss in simulated storage at 40°F and 100°F of fresh and IQF peaches retort pouch-packed by a commercial contractor and peaches repacked at Natick from previous packing in unlined tin cans into pouches. An analysis of the data follows.

Ascorbic Acid Content

Figure 1 shows ascorbic acid concentration at 0, 3 and 6 months for the total pouch contents. The processed fresh peaches contained about four times the ascorbic acid content of the IQF and repacked peaches. The fresh peach ascorbic acid content declined sharply with a decreasing rate, losing about 600 ppm in three months. The other two peach types experienced declines in the first three months but appeared to stabilize at between 3 and 6 months. Except for the repacked peaches at 6 months, there was little evidence of an increase in reducing power, which might come from Maillard sugar-amine reductones and might not be distinguished in this test from ascorbic acid.

Fluorescence Increase

Figures 2, 3, 4 and 5 show the anomalous changes in fluorescence

ratios, absolute emission, and critical wavelengths associated with the fresh fruit during storage. Both fluorescence emission ratio (Fig. 2) and absolute fluorescence emission from excitation, at 395 nm (Fig. 3) show this very clearly. Figure 4, the maximum excitation wavelength from 470 nm emission setting, shows the strong red shift of the processed fresh fruit after 100°F storage. Figure 5, the maximum emission wavelength from 395 nm excitation, also shows the processed fresh peaches changing more dramatically with storage, but in addition displays another diagnostic feature. Fluorescence arising from oxidized ascorbic acid-amine browning, when excited at 395 nm, can be sharply distinguished from sugar-amine browning because the emission is in the 460-480 nm range, whereas that from sugar-amine browning is well above this. We have studied this differential both in in vitro model systems and with heat-stressed fruit to which ascorbic acid has been added. The latter will be reported on below.

Reflectance

The results of reflectance studies are shown in Figure 6, which is a plot of Hunter L value, a measure of lightness versus time in storage. Declines in this value are indicative of browning. Again, the processed fresh sample with the very high ascorbic acid shows the greatest darkening, although the IQF fruit also darkened strongly. The repacked sample, containing a measured 30 ppm tin, is conspicuously lighter initially and changes at a lower rate with time in storage than either IQF or fresh processed peaches.

pH

Figure 7 shows the pH values of the total pouch contents. Although the repacked fruit is noticeably higher in pH, the fresh sample differs little from the IQF in this parameter, and the total spread of the values has been magnified for this figure. The fact that, as we found previously, samples with lower pH can experience more browning and fluorescence, is a cogent argument against sugar-amine browning, for which the opposite behavior is expected.

Sugar Content

Changes in content of fructose, glucose, sucrose and total sugars are shown in Figs. 8, 9, 10, and 11. Hydrolytic production of fructose and glucose and loss of sucrose because of the low pH and high temperature of the simulated storage, proceeded at a high but slowly decreasing rate during the period. The rate was noticeably greater for fresh and IQF than for the repacked fruit, and is in agreement with the relative pH values. Noteworthy, however, is the total sugar content, which, except for the IQF fruit, remains nearly constant. Any

appreciable sugar-amine browning would have markedly reduced the total sugar.

HMF and F Content

Figures 12 and 13 show the changes in 5-hydroxymethylfurfural and furfural content. The major increment of change in concentration of both HMF and F occurs in the interval from 3 to 6 months, not in the interval from 0 to 3 months in contrast to the changes observed from 100°F storage for reflectance, fluorescence or increases in glucose and fructose.

Discussion of Peach Storage Simulation Study

Fortunately, for elucidation of mechanism, the data demonstrated a very large and anomalous initial ascorbic acid content in the fresh peaches compared to the IQF and repacked fruit, and very large commensurate increases in oxidative Maillard browning in the fresh fruit only, as evidenced by fluorescence ratios and absolute emission. Although pronounced acid hydrolysis of sucrose occurred, total sugars changed little, contrary to what would be expected from sugar-amine browning. Furthermore HMF, which is a volatile indicator of sugar-amine browning, was produced in only trace amounts, about 0.01% of the reducing sugars present, which are the most probable precursors. Unless one considers HMF to be a very transient intermediate at a very low steady state concentration, its low level provides little support for much sugar-amine or caramelization loss of sugars, in agreement with the relative constancy of total sugar, above.

As noted above, although the Hunter reflectance lightness factor (L) was most reduced by high temperature storage in the fresh fruit, the IQF fruit also darkened strongly. Because of the relatively low fluorescence levels in the latter, it seems probable that this is residual autooxidative phenolic browning occasioned, in the absence of tin, by the low ascorbic acid.

CONCLUSIONS

As noted above, the occurrence of anomalously high ascorbic acid content in the fresh peaches in simulated storage presented an opportunity to understand the mechanistic role of ascorbate in the browning reactions, as contrasted with that of the IQF and repacked peaches. The concomitant occurrence of very high Maillard indicating fluorescence in the same fresh peaches after storage is strong evidence for the major contribution of oxidized ascorbic acid-amine browning to

the color changes at pH 4 in retort pouch-packed fruit. The strong pro-browning effect of air in the headspace and of softening due to low pH, which was found in earlier studies and was noted above, are strong arguments against a dominating role of the sugar-amine type of Maillard reaction as is the approximate constancy of total sugars. The fact that comminution enhances the browning in pasteurized fruit containing high added ascorbic acid, which is therefore presumably immune to enzymic phenolic browning, also is consistent with oxidized ascorbic acid and not sugar-amine browning, since both sugars and amino acids coexist in the vacuole, whereas ascorbic acid in these fruits is largely an additive.

Together with the evidence from the simulated storage of peaches that the repacked peaches containing 30 ppm tin were the most resistant to the quality changes reported here (although they did suffer more softening than the fresh and IQF fruit), one may conclude that sugar-amine browning is not a major problem in the fruit reported here. One important color and flavor loss problem in retort-pouch packing of wet fruit is protection from enzymic or autoxidative phenolic browning during the peeling, dicing, filling and sealing stages prior to pasteurization and even during storage. Ascorbic acid protects the color very well, but both the thermal process and storage generate oxidized ascorbic acid products which, under the heat stresses mandated by military storage conditions, can cause very detrimental color and odor changes, particularly if there is substantial air in the headspace.

Protective measures against ascorbic acid-amine browning are radically different from those for sugar-amine browning. Proposals to reduce sugar-amine browning by exclusion of reducing sugars are not really possible because some fruits (e.g., pears) contain naturally high levels of fructose, a reducing sugar; also, with time in storage, hydrolysis of sucrose invariably produces two reducing sugars, glucose and fructose. The lowering of pH proposed to decrease sugar-amine browning by prevention of the initial attack of the amino group has been shown³ to be counterproductive when high ascorbate levels are present.

On the basis of the studies reported herein, we propose two effective measures to be taken to prevent oxidized ascorbic acid-amine browning. The first is to keep concentration of ascorbate used to minimize enzymatic browning prior to processing at the minimum required level and the second is to reduce oxygen in headspace as well as total headspace to prevent oxidation of residual ascorbate. Although ascorbate does function as an oxygen scavenger, the subsequent deleterious reactions of oxidized ascorbate are contraindicative of the use of ascorbate as an oxygen scavenger for fruits. Alternative

water-compatible oxygen scavengers such as cysteine, which has been used in pear juice concentrate and apples^{9,10} may be employed, but the simpler expedient may be to use an effective hot fill with steam or nitrogen flushing or a supplemental mechanical vacuum during the pouch sealing prior to thermal processing. Alternative chemical approaches to find functional alternatives to the protective effects of ascorbate without some of the deleterious side effects of oxidized ascorbate will be the focus of future efforts in our laboratory; these approaches include: 1) chemical reduction of both residual oxidized ascorbic acid and quinones by including cysteine (1.5 mM) in the syrup, 2) inclusion of soft tin metal within a removable porous package within the pouch to serve as a reservoir of reducing power and 3) use of a cation exchange resin -- in bead or film form within the package -- to sequester the reactive amino acids.

ACKNOWLEDGEMENTS

The authors express their gratitude to Michael Di'Nunzio for able technical assistance in measurement of sugars and furfurals. Thanks are also due to Dr. K.A. Narayan and Ms. Bonita Atwood for expert advice and editorial review. The indispensable aid of Ms. Ann Marie Huwe in preparing the manuscript is gratefully acknowledged.

REFERENCES

1. Porter, W.L., E.D. Black, C. Anastasia, and K.A. Narayan, "Reduction of Maillard Browning in Wet-Pack Fruit by Amino Acid Sequestration". Technical Report Natick/TR-88/052. U.S. Army Natick Research, Development and Engineering Center, Natick, MA. p. 175 (1988). (AD-A 200 977)
2. Hodge, J.E. and E. M. Osman, "Carbohydrates". In Food Chemistry, Ed. O.R. Fennema, Marcel Dekker, Inc., N.Y., p. 51 (1976).
3. Kluter, R.A., C.P. Dunne, R. Popper, and D. Nattress, "Optimization and Storage Stability of Fruits Processed in Retort Pouches", *Poster Session Paper presented at the 50th Annual Meeting, Institute of Food Technologists, Chicago, Ill., 27 June 1989.
4. Porter, W.L., E.D. Black, Y.-K. Kim, and L. Hoke, "Development of Rapid Methods to Monitor Oxidative and Maillard (Sugar-amine) Polymerization in Energy-Dense, Encapsulated Model Systems." Technical

PORTER, GRANT ET AL.

Report, NATICK/TR-86/063, U.S. Army Natick Research, Development and Engineering Center, Natick, MA, p. 15 (1986) (AD A 179 574)

5. Singleton, V.L. and P. Esau. Phenolic Substances in Grapes and Wine and Their Significance, Acad. Press, NY, pp. 136, 142, 179, 196 (1969).
6. Methods of Vitamin Assay. Ed. J. Augustin, B. Klein, D. Becker, and P. Venugopal, John Wiley and Sons, NY, Chapter 12 (1985).
7. Lee, H.S., R.L. Rouseff and S. Nagy, J. Food Science, 51:1075 (1986).
8. Wucherpfennig, K. and D. Burkardt, Fussiges Obst., 9:416 (1983).
9. Montgomery, M. W., J. Food Science, 48: 951 (1983).
10. Walker, J.R. L. and C.E.S. Reddish, J. Sci. Fd. Agric., 15:902 (1964).

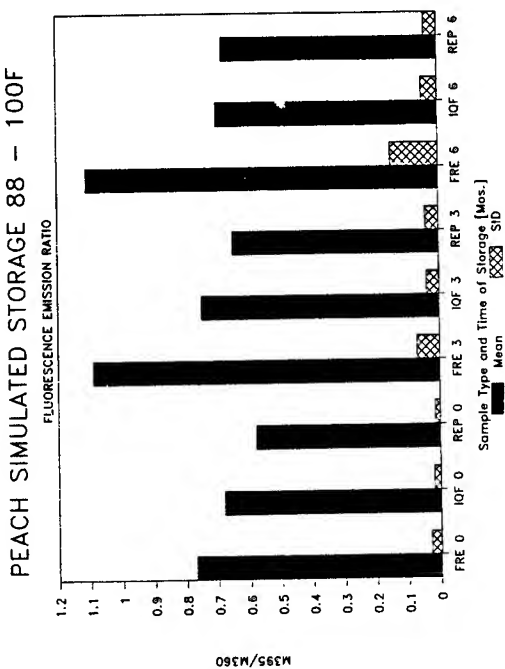


Figure 1. Ascorbic acid concentration.

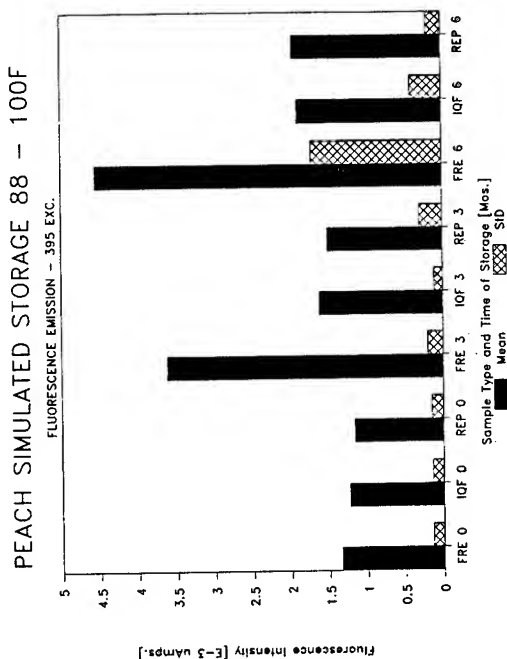


Figure 2. Fluorescence Emission Ratio.

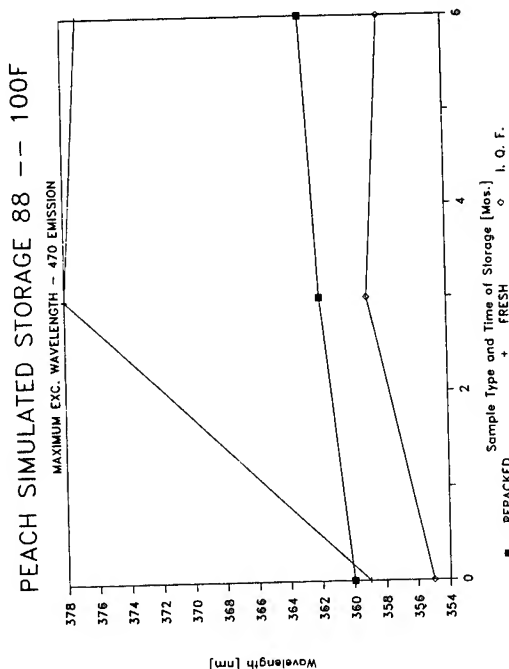


Figure 3. Fluorescence Emission.

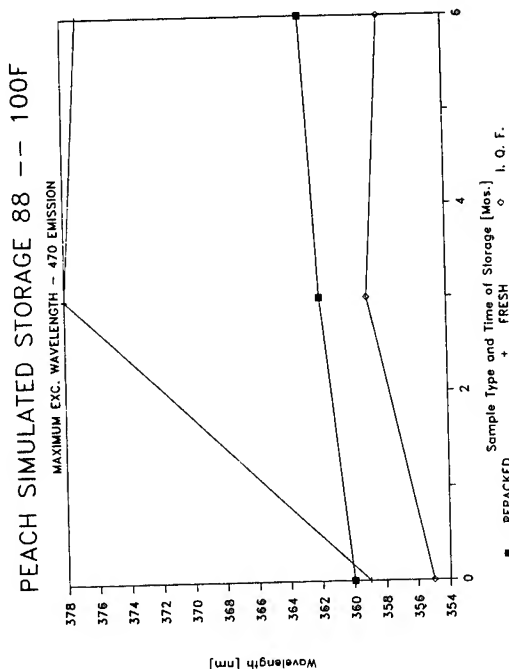


Figure 4. Maximum excitation wavelength.

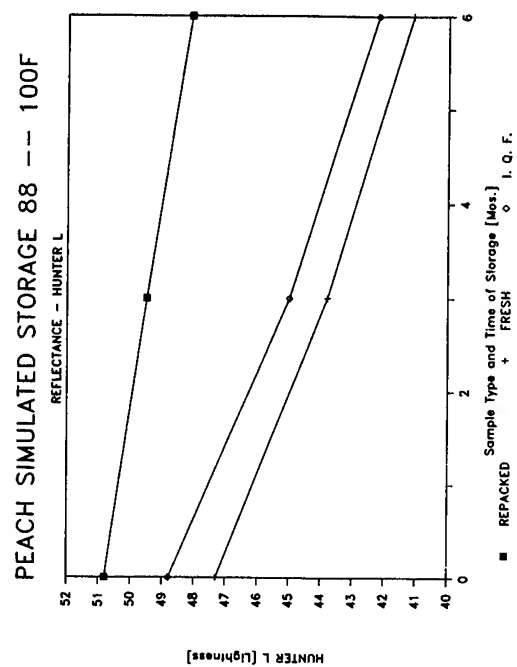


Figure 5. Maximum emission wavelength

PEACH SIMULATED STORAGE 88 --- 100F

Figure 6. Reflectance - Hunter L Value.

PEACH SIMULATED STORAGE 88 --- 100F

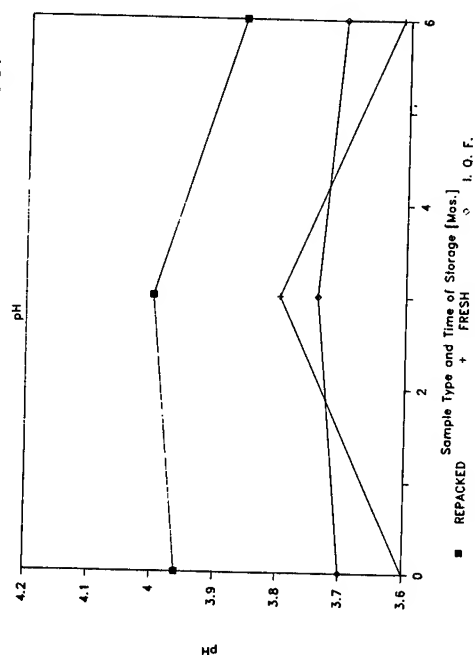


Figure 7. pH of Samples.

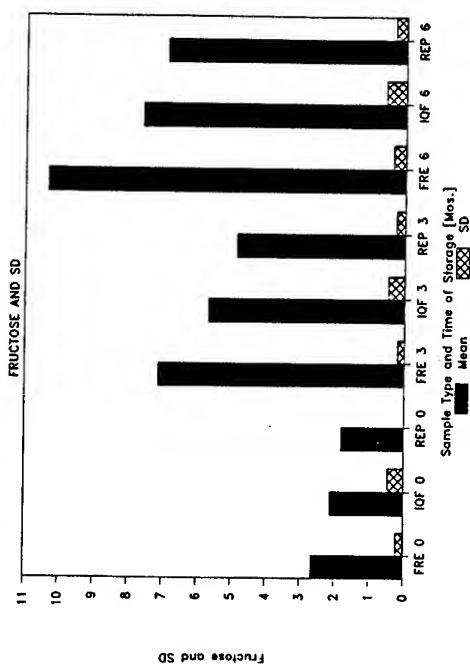


Figure 8. Fructose concentration (g/100g).

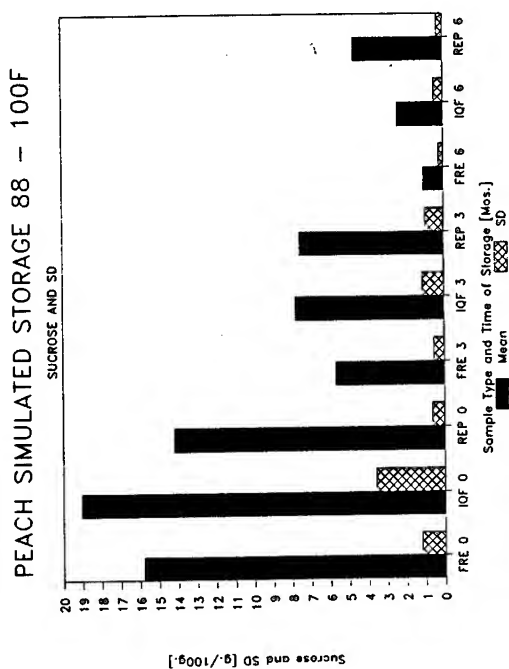


Figure 10. Sucrose concentration.

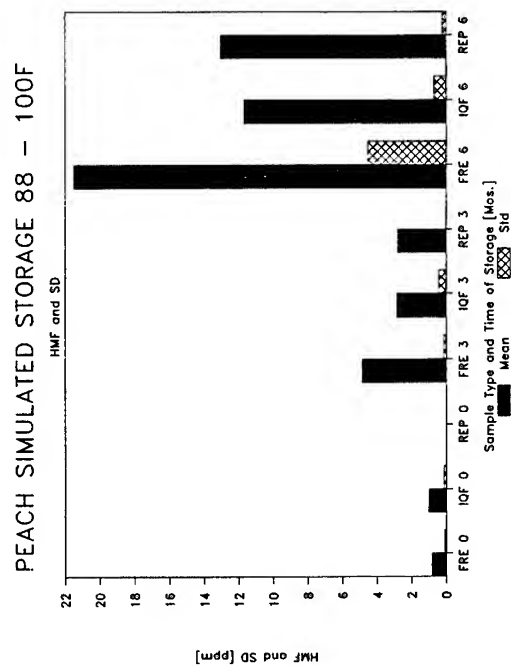


Figure 12. HMF concentration.

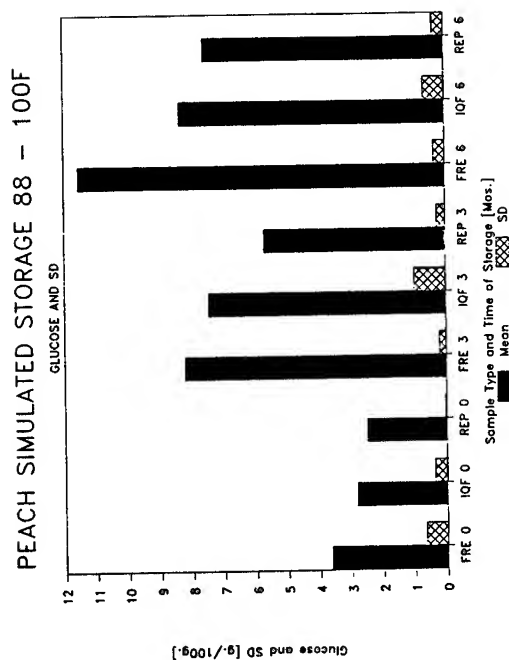


Figure 9. Glucose concentration.

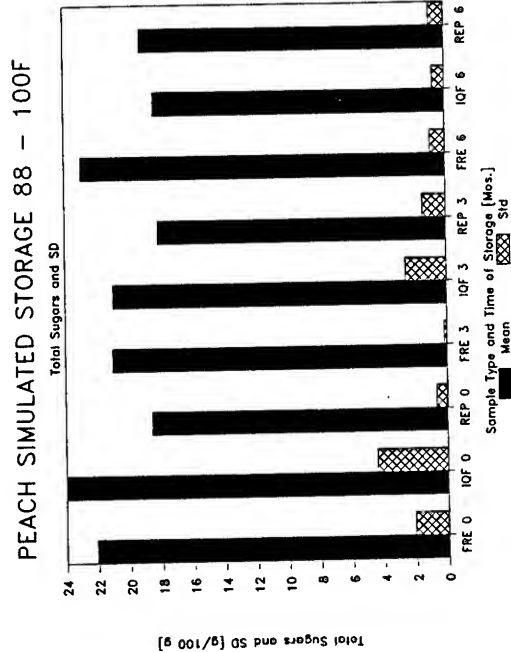


Figure 11. Total sugar concentration.

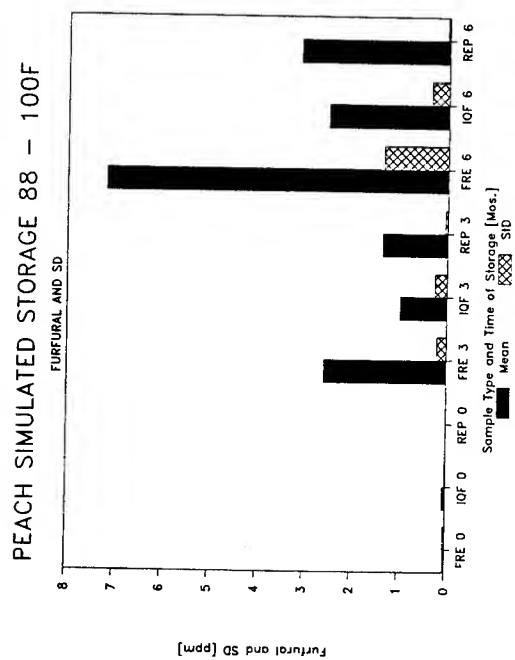


Figure 13. F concentration.

A Layered Architecture for Interfacing Ada and SQL (U)

Judith D. Richardson

US Army CECOM
Center for Software Engineering
Advanced Software Technology
Ft. Monmouth, NJ

Dr. Thomas J. Wheeler

US Army CECOM
Center for Software Engineering
Software Engineering Research Laboratory
Ft. Monmouth, NJ

INTRODUCTION

The method of developing software systems is currently experiencing the largest change since the introduction of general purpose operating systems. Large subsystems conforming to standards are becoming available to be used as the basis for most major software system construction. This increased use of subsystems developed to defacto, national and international standards is forcing the development strategy to become one of composing ("gluing") components together rather than total system development from scratch. The main reason for this standardization trend is the recent leap in the size of software systems and the resulting economic impracticality of developing each system from scratch. This component composition strategy forces developers to address software architecture issues early and requires a standard paradigm for interfacing the application specific software to the standard software subsystems.

Another factor influencing the development strategy is the use of multiple paradigms within a single system. Systems are incorporating substantial subsystems based on specialized paradigms, such as, subsystems utilizing artificial intelligence or database principles. Components developed in these areas are often written in languages other than Ada, for good reasons. This leads to systems written in multiple languages.

Both of these strategic trends lead to a system structuring pattern of interconnected subsystems with some components written in Ada using a modular or object-oriented paradigm and other subsystems written in other languages, such as, LISP, Prolog, C, or SQL, and using completely different paradigms.

This paper develops an interfacing paradigm designed to enhance these component composition and multiparadigm styles of development. The interfacing paradigm is based on a layered approach to interfacing components written in different languages and different styles. The interface bridges the language differences via a syntactic transform and the style differences via a semantic transform. This interface paradigm grew out of a more specific effort to interface Ada and SQL, but is designed with a much broader scope in mind.

The paper will first discuss the Ada/SQL problem to give a context for understanding the paper. It then discusses the issues in general and Ada/SQL issues in particular. It then describes the general two layer architecture approach to interfacing and give some other examples of where the approach has been used. It then returns to the Ada/SQL problem and shows how the interfacing issues are assigned to layers, followed by an outline of a design method for the development of this type of system. The paper concludes with an example showing code from the various layers.

THE ADA-SQL PROBLEM

The standardization on the use of Ada as a programming language and SQL for relational databases implies a need for the development of a standard interface between the two. However, there are incompatibilities in the paradigms of the two languages. Ada is a procedural programming language which takes advantage of strong typing, data abstraction, and complex data structures. SQL, on the other hand, is a query language for relational databases which uses simple data types and flat record structures (or tuples) providing a collection of persistent data. SQL is designed to allow efficient processing of declarative queries on large databases, something that cannot be readily done in Ada. Objects in Ada are accessed by naming the variables while SQL entities are accessed by describing properties of tuples. Ada programs also take advantage of software engineering techniques such as information hiding and modularity, which do not map easily onto SQL. Each language has strengths which should be made available to programmers.

A program using both Ada and SQL is an instance of multiple paradigm programming. A layered interface architecture is one approach to resolve the mismatch which is inherent in programs using two different paradigms. A layered interface attempts to reconcile the differences and at the same time provide access to the features of both languages or environments. In the case of Ada and SQL, this approach provides an Ada program with access to the SQL querying mechanism in order to retrieve and modify persistent information in a database while hiding the syntax and details of SQL.

History :

The project in which this interfacing technique was developed was started at the request of the Ada Joint Program Office (AJPO). They had recognized the need for a standard method of interfacing Ada and SQL and held a workshop to look into approaches [Wheeler 87]. The issues raised at the workshop centered on two areas: the architecture for providing access to database facilities and the binding method for transferring information across the boundary between the two languages. The consensus of the workshop participants was that a modular interface architecture was the "right" approach. Of the methods of binding that were presented at the workshop, the AJPO decided to pursue the module language method. This approach uses the module language specification in the ANSI SQL Standard [ANSI 86]. To establish guidelines for this

binding, the AJPO established the SQL Ada Module Extension-Design Committee (SAME-DC) chaired by the SEI [Graham 89]. The work discussed in this paper grew from involvement in the above efforts.

The SAME effort has developed an approach which handles traditional database application programs in Ada. We, at CECOM, were also interested in exploring the needs of non-traditional database application domains as they relate to the use of Ada and SQL. In particular we were interested in addressing the needs of command and control systems. These systems tend to be large, database oriented programs that provide complex functions to the user. Also, the objects in these application programs are usually more structurally complex than what is found in traditional database applications. The approach of command and control system developers is to use object-oriented development strategies. We wished to form an interface solution that would support systems that use complex objects in an Ada application, but which currently need to store and maintain them in a relational database.

ISSUES

Multiple Paradigms :

A paradigm is a model of a particular thought pattern used to solve a class of problems. A paradigm supplies a systematic, cohesive approach to looking at a problem and thereby assists the solver to think more clearly about the possible solutions. The trade-off is that a paradigm which helps solve one problem may not be flexible enough to solve a wide range of problems [Zave 89]. As a result, large complex systems need to make use of multiple paradigms. Multiparadigm systems are made up of subsystems, each utilizing a different paradigm suited to its particular needs. This allows the overall system to take advantage of the strengths offered by the separate paradigms [Hailpern 86]. The major problems associated with this approach are the mechanism for transferring information and control between the different paradigms and translating between meaningful representations. One approach to solving this problem is to build an interface layer which provides a mapping at the conceptual level of both syntax and semantics. In the case of Ada and SQL this means that systems written in Ada which need the ability to store and retrieve persistent data may desire the ability to utilize a relational database through SQL. To build an interface, the differences between the two languages must be examined.

Data Models :

A data model is a mental model used to abstract the real world into a form that can be used to solve some problem. It contains a representation of real world objects and a mechanism for representing relationships between them as well as a set of operations for their manipulation. Each data model gives a "flavor" to the solutions derived from its

use. It is through the use of a data model representation that approaches to problem solving are supported.

The relational model was first proposed by Codd [Codd 70]. In the model the abstract representation of data is a collection of tables, which are treated as mathematical relations, made up of named columns called attributes, with each row being called a tuple. In the relational model values are obtained by describing the attributes of the tuple, and not by the path to its location, which was the mechanism used in the earlier network and hierarchical models.

The entity-relationship (ER) model was proposed by Chen to form a more generalized data model [Chen 72]. The ER model was an attempt to form a unified view of data. The ER model takes a fundamental view of the data and so is a useful tool in developing the conceptual model of a database. Although it is most often used as a first step in developing a relational database, it can be used to form the initial analysis for other types of databases.

In an object-oriented database, entities are objects of abstract data types. This means they need not be atomic, as in the relational model, but may be complex structures. Associated with an object are classes (types) and messages (operations). Every object is derived from some class. Classes, in turn, may be derived from other classes forming a classification hierarchy, which allows inheritance. Objects are encapsulated, hiding their structure. Access to the information in them is provided through the messages defined on their class. Named relationships are supported between objects. They form navigational links, as in the network or hierarchical databases, and predicate based links, as in relational databases [Zdonik and Maier 90].

The data model for Ada uses non-persistent, strongly typed, named objects (~entities). It can support objects as abstract data types. Abstract data types are formed using records and packages, which contain an arbitrary collection of types. The fields of a record contain the object's information (~attributes). Access to an object (~relationships) can be direct, through the object's name, or navigational, through access pointers. One of the strengths of Ada's data model is its flexibility; its ability to support a variety of problem solving approaches. The Ada/SQL interface should not limit the application program to the relational data model, therefore it needs the ability to form an abstraction of the services provided by the database.

Views :

Views are a mechanism used to partition a database in order to form abstractions or hide information not needed or accessible to a user or group of users. For example, in a company's database each department would have their view of the database, i.e., the accounting department's view, the personnel department's view, etc. Each view reflects the users model of the world. Views are not stored in the database, only the definition of the view. Application programs are written against defined views. Since the view provides

modularization of the access to the database from the application program, the view definitions form the modularization used in the Ada/SQL interface.

Typing :

The typing model in SQL is limited to a set of basic data types and a dynamic collection of tuples of those types. The ANSI standard defines eight data types, although different implementations provide slightly different interpretations. Ada on the other hand is a strongly and statically typed language. The language provides the user with the ability to define new types and subtypes at compile time, as well as structures. These features ensure that operations only occur on the intended type. Along with strong typing, Ada provides for data abstraction. Data abstraction is the defining of a type by specifying the operations that apply to the type while hiding the implementation from the user. Any interface solution must provide a mechanism for converting between the primitive types and tuples in the database, and the user defined types in the application program. Additionally, the interface should check that only valid data is inserted or modified in the database, therefore consistency constraints can be assured for any given type.

Structure of Entities :

The structure of entities in relational databases is first normal form. This means that all the attributes of an entity are atomic. Therefore, they do not have subparts and can be viewed as simple records of primitive types. Relationships between entities are just shared, named fields which allows linking mechanism, via joins, between different records. In traditional database applications this is sufficient, however there are application domains where structured entities are needed. One domain is CAD/CAM systems, another is programming environments and a third is Command and Control systems. An Ada program can easily support the structures used in these various application domains. However, they cannot be derived by simple, direct mapping from SQL tuples.

Mapping Models:

If a one-to-one mapping is not possible between two representations then there is a structure clash [Jackson 75]. To map between the structures, a process for transformation is developed. First the lowest common denominator between the two structures is determined, then algorithms are developed to decompose the first structure and then to build the second structures up from primitives. But, whereas Jackson allows the data structures to be visible to the users, Parnas hides as much of the details as possible [Parnas 72]. Parnas decomposes a system into modules based upon the concept of encapsulating design decisions into modules and then supplies the user with an interface to the module that only provides the information needed and nothing more. This technique of information hiding isolates the details of the mapping process inside of the interfacing module when there is structure clash.

Abstract Interfaces :

Parnas also describes the use of abstract interfaces [Parnas 77]. By abstract, Parnas is referring to the representation of several actual objects without specifics to any particular object. The essence of the real world is captured in the abstraction which can then be re-used to other objects that the abstraction represents. The interface is the format of the information exchange between two software components along with all the assumptions that each makes about the other. An abstract interface is defined as a set of assumptions that represent more than one possible interface. An abstract interface models those properties that the interfaces hold in common and hides the differences.

LAYERED INTERFACE ARCHITECTURE

Two Layer Interfaces :

One way to map the mismatch between multiple paradigms is to build a two layer abstract interface. The upper level interface provides the actual interface to the user, which reflects the conceptual model appropriate to the application program, thus the application an appropriate model reflected in its interface. It uses the concepts of data abstraction to generalize the information crossing the interface, abstracting from particular characteristics of the external subsystem. The lower level interface reflects the semantics of the external system to the model transformation in the two layer interface, hiding the details of the external subsystem.

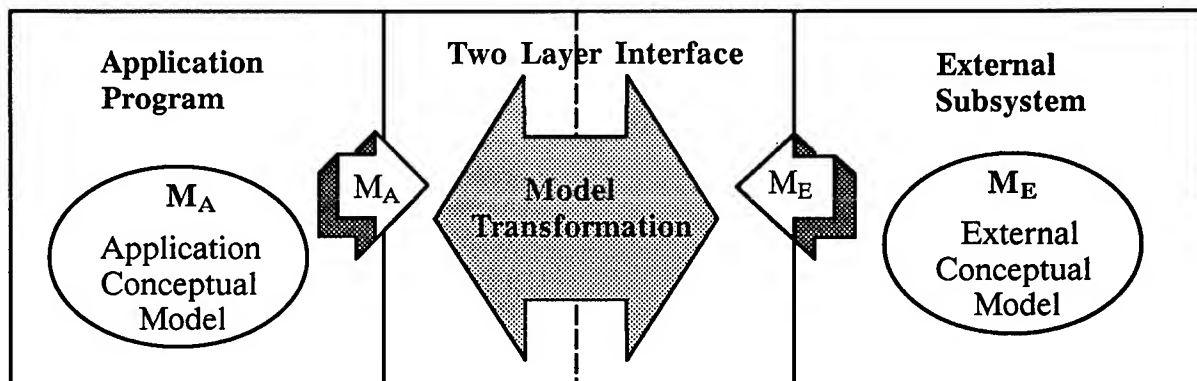


Figure 1

In addition to the semantic transformation provided by the abstract interface, there is a syntactic (language) transformation which must occur if the languages, in which the application and external subsystems are implemented, are different. This syntactic trans-

formation may occur anywhere within the abstract interface. The syntactic transformation is often called the binding.

For example, in UNIX™, when a user requests a file be copied to the current directory, the command `cp` along with the path to the file are all that the user has to supply. The operating system determines from the path which device the file resides on and makes the call to the appropriate device driver. The device driver is concerned (if the file were on disk) with such things as: seek algorithms, inodes, motor speed, etc. The supplemental information and work involved in copying the file are transparent to the user. The upper level provides the file abstraction, the lower level the device read/write.

Another example is a proposed two level architecture for communication within multi-Ada program systems [CASS 89]. The interprogram interface is defined in terms of remote procedure calls (RPC). In other words, to the application program the location of the called procedure is unknown. Again, the lower level is concerned with the details of the communications system, such as, sockets and communication protocols. The upper level provides an abstract interface to the user, taking care of the bundling of the message with logical source and destination names. The upper layer provides a procedural interface defining services used by the application, the lower layer a message passing transport service.

A third example is the interface between the real world and a control system. In a control system certain observable characteristics of the external world are monitored by sensor subsystems which understand the meaning of those characteristics. In this example the lower layer of the interface receives the sensor data coming in from the external world while the upper layer interprets the information into a meaningful representation. An example of this interface is a stop light control system [Wheeler 86]. The information from the sensors is passed to the lower layer of the interface. It is then sent to the upper layer which converts the information to a representation of traffic presence or absence in the various lanes. The upper level is concerned with control of the intersection, the lower level with the sensors and relays.

As can be seen in each of these examples, the upper layer presents a model which is meaningful to the application, while the lower layer uses a model that is meaningful to the external subsystem.

ADA AND SQL INTERFACE ARCHITECTURE

Architecture

Applying this architecture to the Ada/SQL Interface, resulted in a layering whereby the upper layer provided the semantic transformations, becoming the Semantic Layer, while the lower layer provided the syntactic transformations, becoming the Syntactic Layer (Figure 2).

The Syntactic Layer contains the binding of SQL to Ada. It interprets the SQL error codes and takes care of null values. In general, it takes care of the implementation details associated with the chosen binding and hides them from the application program. This layer directly maps the types in the database to predefined Ada types. The procedure calls to the Syntactic Layer reflect the syntax of SQL. The implementation that we use of the Syntactic Layer is based upon the Module Language definition in Chapter 7 of the ANSI SQL Standard [ANSI 86].

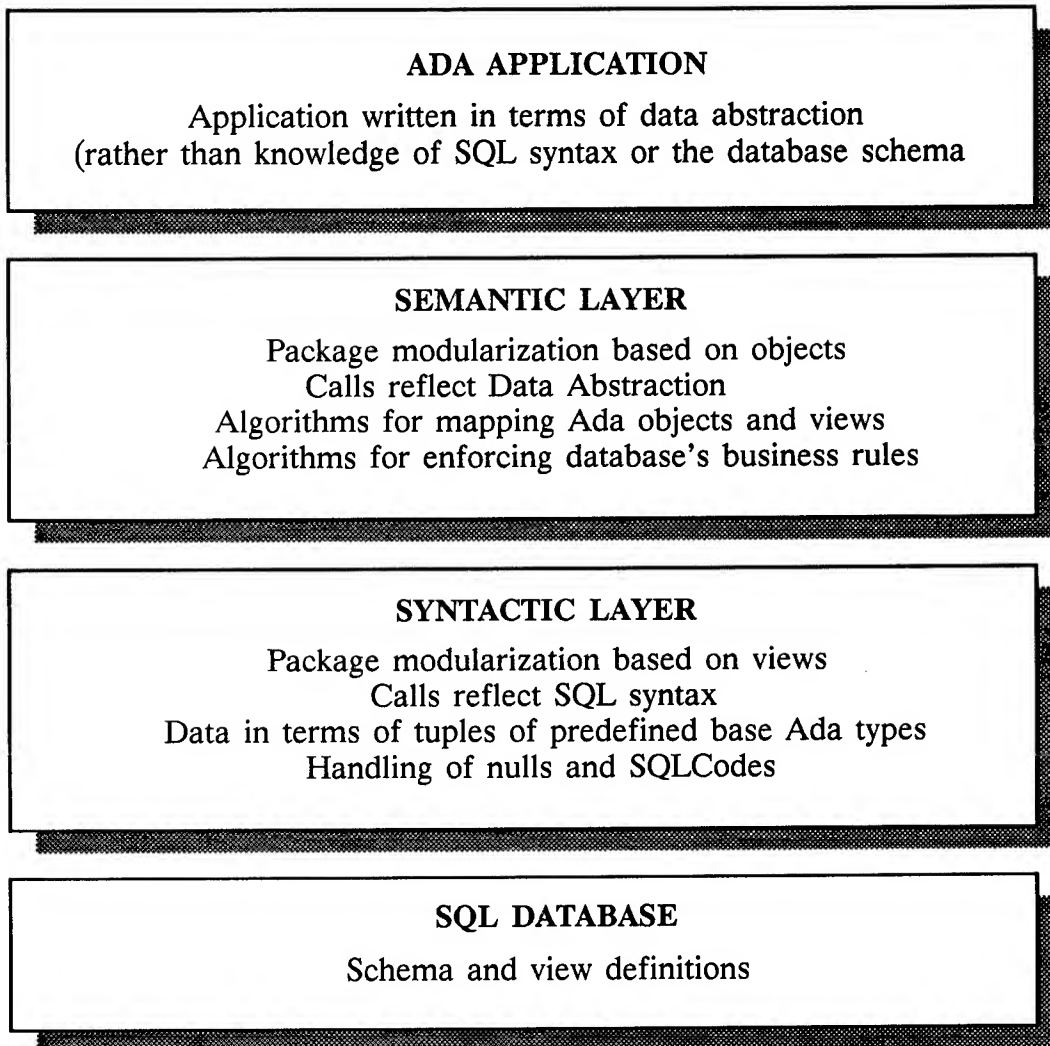


Figure 2
Interface Architecture

The Semantic Layer is the visible part of the interface to the Ada application programmer. This layer converts the tuples from the Syntactic Layer to data structures forming Ada abstract types. The data abstractions reflect the information needs of the Ada application program and are derived from view definitions in the database. This layer not only provides the application with access to the information contained within the data abstraction, but also maintains data integrity by enforcing the business rules governing modification of the database. It is the job of the interface to enforce the rules and modify the original tables giving the appearance of modification through views even when that is not supported by the DBMS. Calls to the Semantic Layer abstract away the SQL functions used in the lower level. The upper level of the interface can be viewed as a retriever and maintainer of a persistent abstract data type. This design takes advantage of Ada's ability to implement information hiding and data abstraction.

Use of this architecture gives Ada programmers the advantages of an SQL DBMS without requiring them to be proficient in SQL or understand the logical schema of the database. This approach not only segregates the two paradigms from a design standpoint, but also has practical advantages, since only the builders and maintainers of the interface need to know both worlds.

Design Method :

Designing a two layer interface for Ada and SQL, using the architecture described, consists of three phases: initial analysis, mapping of views and data abstractions, and business rule enforcement (Figure 3). The first step is a general analysis of the application in order to develop a conceptual model of the problem. This can be accomplished using the ER model to form an object based view of the problem domain. By using the ER model to describe the objects, there need not be a presumption toward a relational solution. At this point a functional analysis is done of the identified objects and the flow of information between objects. Once this initial analysis is completed the design of the interface can begin.

The development of the interface is accomplished by starting at each end and developing the details within each layer. The first stage is concerned with mapping the objects in the database to the objects in the Ada application. To develop objects in the database, a logical model is established from the conceptual model and developed into a normalized relational schema. On the Semantic Layer side, abstract data types and their underlying data structures are developed from the objects defined in the general analysis. Views are then derived from the normalized schema and the information needs of the data abstractions.

The modularization within the interface is based upon the views and data abstractions. At the Syntactic Layer a package is developed for each view. The syntactic packages provide open, fetch, and close operations on the tuples needed by the Semantic

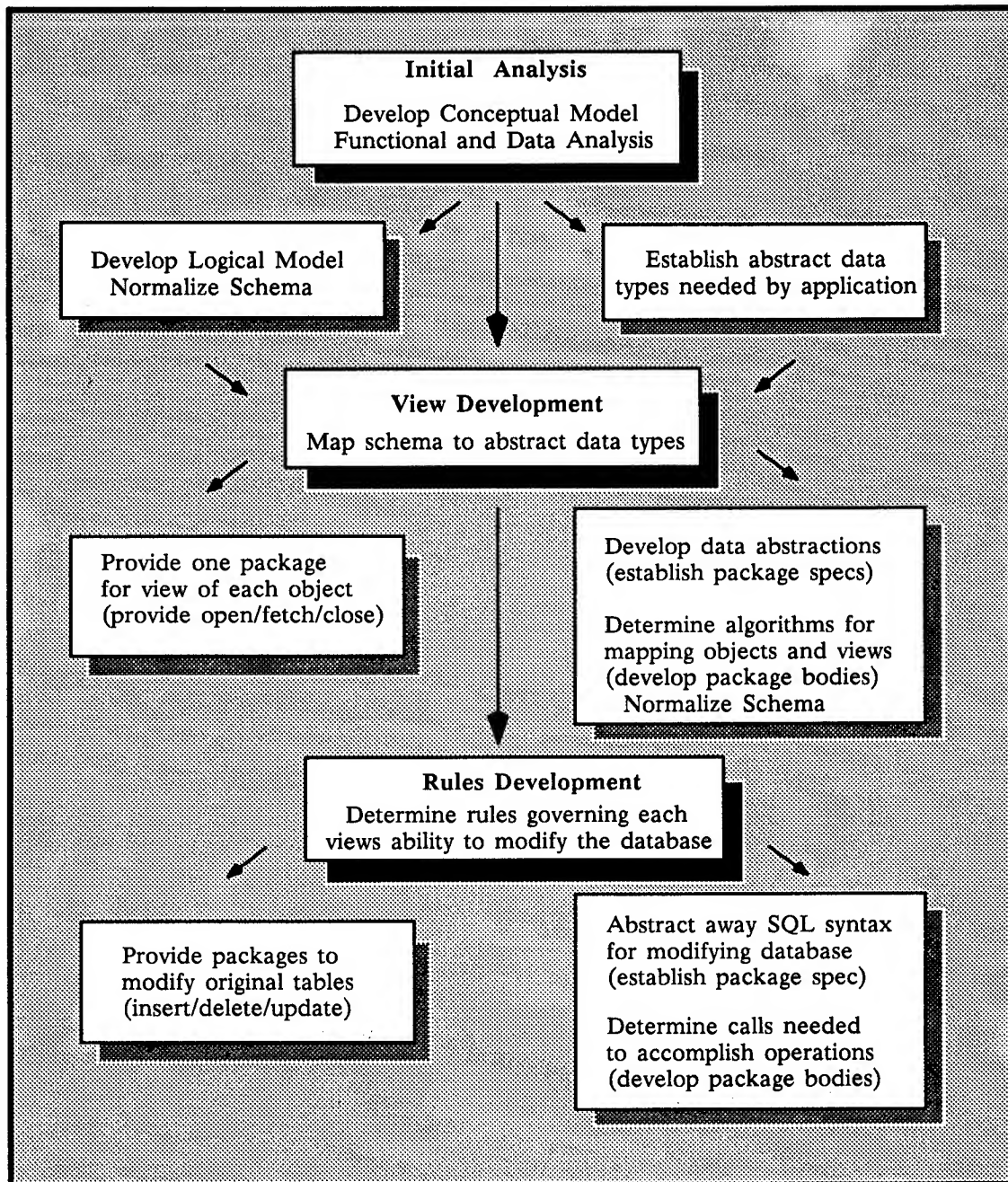


Figure 3
Design Method

Layer. The Semantic Layer develops packages for each data abstraction. The specifications of these packages define the types and logical operations for the types. The syntax of these operations reflect the point of view of the application not the syntax of SQL. The bodies of the Semantic Layer packages define the underlying data structure and the algorithms for building them from the tuples supplied by the Syntactic Layer.

Once the data abstractions are defined and the support for establishing them is complete, the next step is to determine the rules or policies governing modification of the database in order to assure consistency. The modifications in SQL terms can take the form of inserts, updates, or deletes. Since the modification in most SQL DBMSs occur on the original tables not the views, the Syntactic Layer provides packages separate from the view packages for these calls. The Semantic Layer encapsulates the sequence of calls needed to modify the database according to the established rules. Within the Semantic Layer, as in the definitions of objects, the modification requests from the application are in terms of the data abstraction. By hiding the business rules within the interface, systems do not have to rely upon the applications to remember and enforce the rules.

EXAMPLE

As an example we developed a extended ER model of the objects in a motor pool (Figure 4). The relationships between Companies, Motor Pool, and individual Vehicles are fairly straight forward. However, the relationship between a Vehicle and the parts that make up the Vehicle required a representation of the idea of variety (truck or jeep) and made of (vehicles made of multiple, interrelated parts). We wanted to capture in the Vehicle_Record the concept of relative position of a part within the assembled vehicle. In the figure, the white boxes, ovals, and dark lines represent the ER diagram. The ovals indicate attributes of the relationship. They represent the two ways the part-vehicle relationship can be thought in terms of; as a linear list of parts and as parts in a tree structure representing relative position. The two shades of grey indicate the sub-assembly structural information we wished to capture.

The example was implemented on an AT clone running MS-DOS and using the ALSYS Ada compiler, XDB SQL engine, and Computer Associates' Module Compiler to provide the binding. Once it was running under this configuration, it was ported to the Army Tactical Command and Control System (ATCCS) configuration consisting of the Hewlett Packard 9000 Series 300 running UNIX System V and using HP's Ada compiler (derived from ALSYS), HPSQL DBEnvironment, and HP's AdaSQL binding. Even though the bindings were entirely different for these two systems, the only changes made between the two versions were in the bodies of the Syntactic Layer, where the implementation details of a binding are handled.

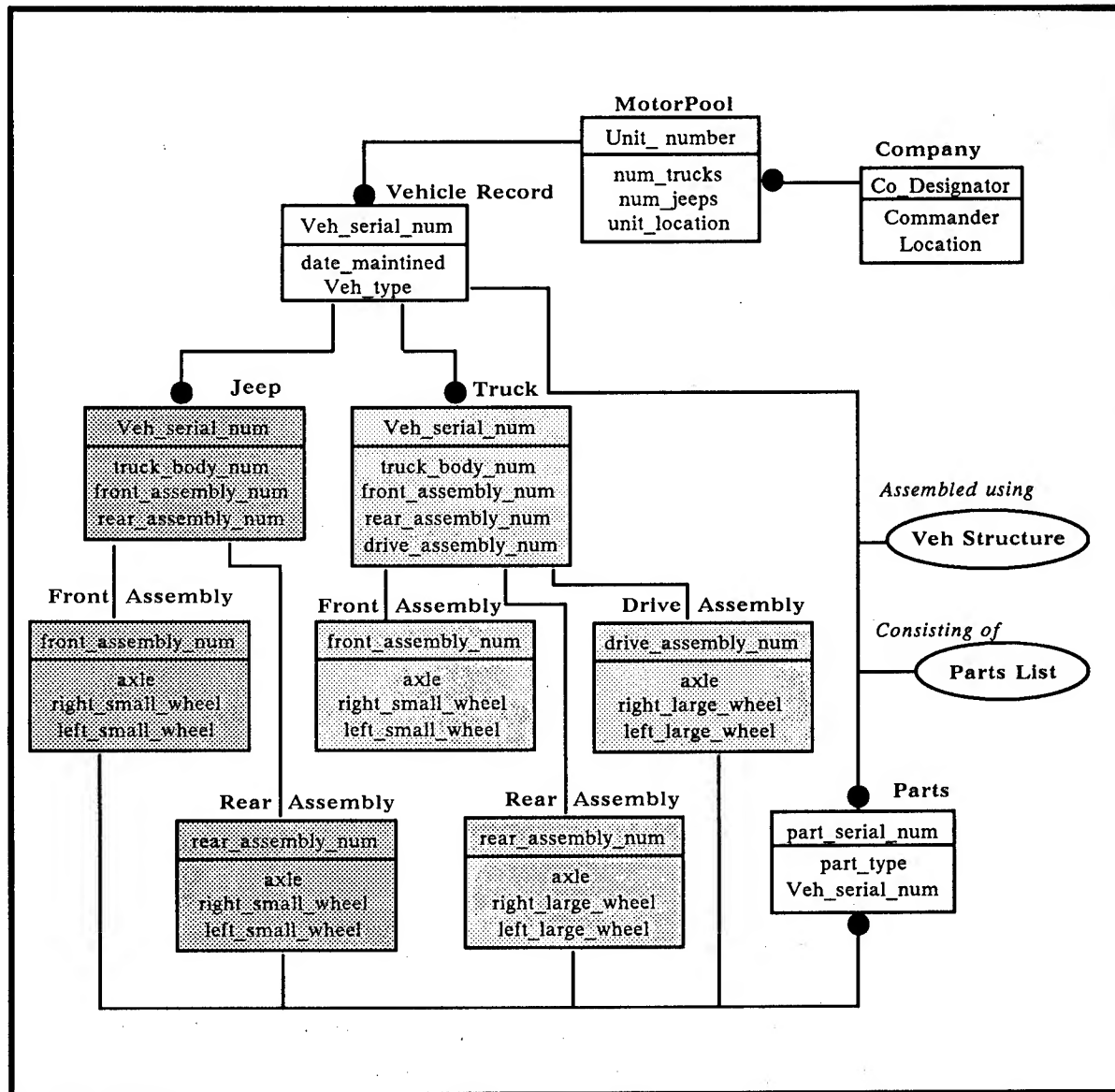


Figure 4
Extended Entity Relationship

An Example of Mapping Views to Data Abstractions :

The following mapping was developed from the normalized relational schema to the objects identified in the extended ER model. This was followed by establishing the type definitions to be used in the bodies of the Semantic Layer. From this the view definitions were developed.

Vehicle_Request

View Definition :

```
CREATE VIEW Vehicles_View
SELECT unit_number, num_trk, num_trn
FROM Motor_Pool ;
```

Fetch :

```
FETCH SELECT unit_number, num_trk, num_trn
FROM Vehicles_View
WHERE unit_number = : Unit_of_Interest ;
```

Syntactic Layer Tuple :

```
type Vehicle_Tuple_Type is
record
    Unit_Number : SQL_Standard.Int ;
    Number_Trucks : SQL_Standard.Int ;
    Number_Jeeps : SQL_Standard.Int ;
end record ;
```

Semantic Layer Object :

```
type Vehicle_Variety_Type is ( TRUCK, JEEP ) ;
type Vehicles_Available_Array_Type
is array ( Vehicle_Variety_Type ) of Number_Of_Vehicles_Type;
type Vehicles_Available_Record_Type is
record
    Unit_Number : Unit_Number_Type ;
    Number_Available : Vehicles_Available_Array_Type ;
end record ;
```

Semantic Layer Function :

```
procedure CREATE_VEHICLE_AVAILABLE
( UNIT_OF_INTEREST : in Unit_Number_Type ) ;
function GET_NUMBER_AVAILABLE ( VEHICLE : in Vehicle_Variety_Type )
return Number_Of_Vehicles_Type ;
```

CONCLUSION

This paper examined the issues in trying to resolve the impedance mismatch between Ada and external systems. It described a two layer interface architecture to deal with both the syntactic and semantic issues resulting from using multiple paradigms. It used the interface of Ada and SQL as an example to illustrate the interface. Use of the architecture gives an Ada program access to external systems in a way that hides the details of those systems and maintains independence of its evolution. This allows the use of commercial off-the-shelf (COTS) software subsystems while maintaining control over the systems architecture and the system's evolution.

REFERENCES

- [ANSI 86] ANSI, *Database Language-SQL*, ANSI X3.135-1986, American National Standards Institute, Inc., New York, 1986.
- [CASS 89] Common ATCCS Support Software Working Group-Architecture Subcommittee, *Requirements for Inter-Task Communication* (Draft), 1989.
- [Chen 72] Chen, P., "The Entity-Relationship Model-Toward a Unified View of Data", *ACM Transactions on Database Systems*, March 1976.
- [Codd 70] Codd, E. F., "A relational model of data for large shared data banks", *Communications of the ACM*, June 1970.
- [Graham 89] Graham, M. H., *Guidelines for the Use of the SAME*, Technical Report, Software Engineering Institute, Pittsburgh, PA, 1989.
- [Hailpern 86] Hailpern, B., "Multiparadigm Languages and Environments", *IEEE Software*, January 1986.
- [Jackson 75] Jackson, M., *Principles of Program Design*, Academic Press, New York, 1976.
- [Korth and Silberschatz 86] Korth, H. F. and A. Silberschatz, *Database System Concepts*, McGraw-Hill, New York, 1986.
- [Parnas 72] Parnas, D. L., "On the Criteria to Be Used in Decomposing Systems into Modules", *Communications of the ACM*, December 1972.
- [Parnas 77] Parnas, D. L., "Use of Abstract Interfaces in the Development of Software for Embedded Computer Systems", *NRL Report*, Number 8047, June 1977.
- [Wheeler 86] Wheeler, T. J., "An Example of the Developer's Documentation for an Embedded Computer System Written in Ada", *Ada Letters*, November-December 1986 and January-February 1987.

RICHARDSON & WHEELER

[Wheeler 87] Wheeler, T. J., *Memorandum for Record*, CECOM Software Technology Division, Information Processing Directorate, November 1987.

[Zave 89] Zave, P., "A Compositional Approach to Multiparadigm Programming", *IEEE Software*, September 1989.

[Zdonik and Maier 89] Zdonik, S. B. and D. Maier, *Readings in Object-Oriented Database Systems*, Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1990.

ACKNOWLEDGMENTS

The authors wish to thank Arturo Ferrer, Rennetta McGhee and Hoang D. Lee for their efforts in developing the example problem.

Vehicle Motion Resistance Due to Snow

Paul W. Richmond III

U.S. Army Cold Regions Research and Engineering Laboratory
Hanover, New Hampshire 05755-1290

INTRODUCTION

Vehicle mobility is dependent on vehicle traction and the resistance due to deformation of the terrain in front of a wheel or track. A vehicle is "mobile" if traction is greater than the resistance. A mobility model requires accurate prediction of both traction and resistance.

During the winters of 1988 and 1989 a winter mobility study program was initiated and jointly conducted by the Waterways Experiment Station (WES) and CRREL. These studies, part of the larger U.S. Army Wheels versus Tracks Program, were to be used to compare predictions of the CRREL shallow snow mobility model with actual snow mobility data for a wide variety of vehicles. The results of the winter studies are reported in Blaisdell et al.,¹ Green and Blaisdell² and Richmond et al.³ This paper presents the mobility resistance data obtained in undisturbed snow and an analysis and development of empirical resistance equations. These equations are based on a parameter that includes both vehicle characteristics and snow properties.

FIELD EXPERIMENTS

The data presented in this paper were obtained during field experiments carried out at the Keweenaw Research Center (KRC) located at the Houghton County Airpark, Michigan (KRC is located on the Keweenaw Peninsula, in northern Michigan). Mobility tests were conducted by personnel from WES, KRC and CRREL. Two types of mobility tests were conducted; traction tests in various snow conditions and resistance tests in undisturbed snow; this paper is concerned solely with the resistance tests.

CRREL and KRC personnel conducted mobility tests using the CRREL Instrumented Vehicle (CIV). The CIV, which is fully described by Blaisdell,⁴ is a 1977 Jeep Cherokee with modifications to its braking and driving components to accommodate typical mobility tests.

In addition to its on-board computer based data acquisition system, the vehicle contains a number of transducers for force and speed measurements. In these and previous tests, the CIV produced results (traction data) that agreed very well with data obtained from larger vehicles with similar ground contact pressures (e.g., Blaisdell et al.¹) in side-by-side tests. For these reasons, and because of its versatility, the CIV was used to extend this winter mobility data base. The CIV was equipped with several different tires during this study to examine the effects of tire parameters on resistance. These tires and their characteristics are shown in Table 1. Two standard all-season mud and snow

Table 1. CIV tire data.

<i>Tire nomenclature</i>	<i>Code^a</i>	<i>Inflation pressure (kPa)</i>	<i>Radius^b (m)</i>	<i>Width^c (m)</i>
Michelin LT235×CH4	A	179	0.375	0.260
		103	0.375	0.279
Firestone T145/80	C	14	0.325	0.156
Goodyear Eagle P225/60R15	D	248	0.35	0.274
Goodyear Tiempo P225/75R15	E	79	0.356	0.254
		103	0.349	0.267

a. Tire code used in Figure 4.

b. Radius of undeformed tire.

c. Maximum deformed width.

tires (the Goodyear Tiempo in 1988, and the Michelin XCH4 in 1989) were selected for their good all-round traction characteristics. Besides these standard tires, two other tires were used during the resistance tests in 1989; these tires were not used for traction tests and were chosen solely based on their width and availability. The Firestone T145/80 is a temporary spare tire with a maximum width of 0.156 m (6.1 in.) and the Goodyear Eagle P225/60R15 is a low profile "street-rod" tire with a maximum width of 0.274 m (10.8 in.).

Personnel from WES conducted resistance tests using the following group of military vehicles:

- M988 High-Mobility Multipurpose Wheeled Vehicle, 4×4, (HMMWV) equipped with Michelin 37×12.5 R16.5LT tires.
- M977 Heavy Expanded Mobility Tactical truck, 8×8, (HEMTT) equipped with Michelin 16.00 R20 tires.
- M973 Small Unit Support Vehicle (SUSV), articulated, tracked
- M2 Bradley Fighting Vehicle System (BFVS), tracked.
- LAV 25 Light Armored Vehicle (LAV), 8×8, equipped with Michelin 12.50 R20 or 11.00 R16 tires.
- M923 5-ton truck, 6×6, equipped with Goodyear 14.00 R20XL tires.
- M113A1 Armored Personnel Carrier, tracked.

These vehicles were chosen to span the full range of typical military ground vehicle contact

Table 2. Tire and track data for military vehicles.

<i>Vehicle</i>	<i>Tire nomenclature</i>	<i>Inflation pressure (kPa)</i>	<i>Radius^a (m)</i>	<i>Width^b (m)</i>
HMMWV	37.00×12.5 R16.5	138/152	0.429	0.33
	36.00×12.5 LT	138		0.425
HEMTT	16.0 R20	241/276	0.617	0.475
		139/207	0.589	0.483
LAV25	12.50/75 R20XL	207	0.445	0.378
		103	0.414	0.343
	11.00 R16×L	290	0.434	0.314
		165	0.417	0.332
SUSV	Track	—	32.5°	0.6096
BVFS	Track	—	25°	0.533
M113A1	Track	—	21°	0.381

a. Radius for wheeled vehicles, entrance angle for tracked vehicles.

b. Maximum deformed width.

pressures and represent a cross section of current military wheeled and tracked vehicles. The characteristics of these vehicles are given in Table 2.

TEST PROCEDURES

The test procedures followed typical mobility field studies, in that, measurements of net traction T_n and total motion resistance R_t were made with each vehicle under varying snow conditions. Although it was desired to conduct tests using all the vehicles on each day, the shortness of the available daylight and the number of vehicles, combined with lack of appropriate snow falls, precluded this. All of the tests were done in areas that had a packed snow base, with the exception of some traction and resistance tests conducted on snow-covered ice (the KRC ice rink). As long as the underlying layer is nondeformable, it has little effect on resistance.

CRREL Instrumented Vehicle

Motion resistance is measured with the CIV having its the rear tires driving and its front wheels rolling free. Since the triaxial load cells are located just inside the front wheels, this test measures the total amount of resistance at the front tires only. Motion resistance is first established on a level, undeformable surface. Measurements of this hard surface resistance (R_h) are obtained for each tire type and selected inflation pressure and subsequently subtract-

ed from all the in-snow resistance measurements. Motion resistance tests are performed at a vehicle speed of 8 km/hr (5 mph). Resistance values are independent of moderate variations (± 2 mph) in vehicle speed. Variations in R_h values between tire types are the result of differences in the forces caused by tire flexing and can be attributed to design differences.

The external vehicle motion resistance in a snow cover (R_s) is calculated by measuring the motion resistance (R_t) in the test area using the above procedure and subtracting the hard surface motion resistance (R_h). The left and right external resistances are calculated separately and the average is used for analysis.

Military vehicles

The motion resistance of the military vehicles was determined by measuring the vehicle's resistance to towing (TMR). Here,

$$\text{TMR} = R_t n \quad (1)$$

where n is the number of tires or tracks and R_t is the total motion resistance on an average tire or track.

The procedure used for obtaining TMR of the test vehicle in each test area was to tow the vehicle (with its transmission in neutral) at a speed of approximately 3.2 km/h (2 mph). After each traction test, the vehicle was steered into an undisturbed area adjacent to the traction test area, usually in a position straddling the ruts of an associated traction test. The test vehicle was then towed forward to determine TMR. The test proceeded for a sufficient distance to permit the load cell (installed on the tow cable) readings to stabilize and to be recorded on magnetic tape. An average value during the stable portion of the record was taken as the TMR, and R_t was then calculated from eq 1. Finally, external resistance due to snow compaction R_s was found by subtracting the average R_h from R_t .

Measurements of snow characteristics (depth, density, temperature and vehicle sinkage) were obtained at each test area while tests were being conducted. The undisturbed snow density ranged from 70 to 320 kg/m³, and depths from 3 to 30 cm. Generally, the snow was from one storm and thus relatively uniform. Air temperatures were always well below freezing.

RESULTS AND ANALYSIS

Sinkage

The amount a vehicle sinks into the snow (sinkage), or the depth of the rut left by vehicle passage, greatly affects the amount of motion resistance exerted by the snow. External resistance (R_s) is strongly related to the amount a vehicle sinks in the snow, and the CRREL shallow snow mobility model predicts external motion resistance based on sinkage. The model

uses the following equation to estimate sinkage (z):

$$z = h \left(1 - \frac{\rho_o}{\rho_f} \right) \quad (2)$$

where h is depth of undisturbed snow (m), ρ_f is the theoretical final snow density (kg/m^3) of the snow, in the rut caused by vehicle passage, and ρ_o is the initial undisturbed snow density (kg/m^3). Table 3 is used to obtain the theoretical final density for the snow, from nominal vehicle contact pressure (tire pressure).

Snow properties, rut depth and densities were measured and compared with sinkage calculated using eq 2 and Table 3. Table 4 presents the measured snow and rut characteristics and the predicted values for two representative vehicles, the HMMWV ($\rho_f = 500$) and the SUSV ($\rho_f = 500$). In general, the measured rut density values do not agree very well with the theoretical values obtained from Table 3. In nearly all cases, the Table 3 final density values are greater than what was measured. However, the sinkage values calculated using eq 2 are similar to the measured values. These and sinkage data for the other vehicles are compared in Figure 1. Most of the data are within 25% of the measured sinkage. Generally, sinkage is under-

Table 3. Final theoretical densities used in the CRREL shallow snow mobility model.

Maximum ground pressure (kPa)	Theoretical final density (kg/m^3)
< 210 kPa	500
211–350	550
351–700	600
> 701	650

Equation 2 was derived by assuming that the width of the deformed snow is equal to the

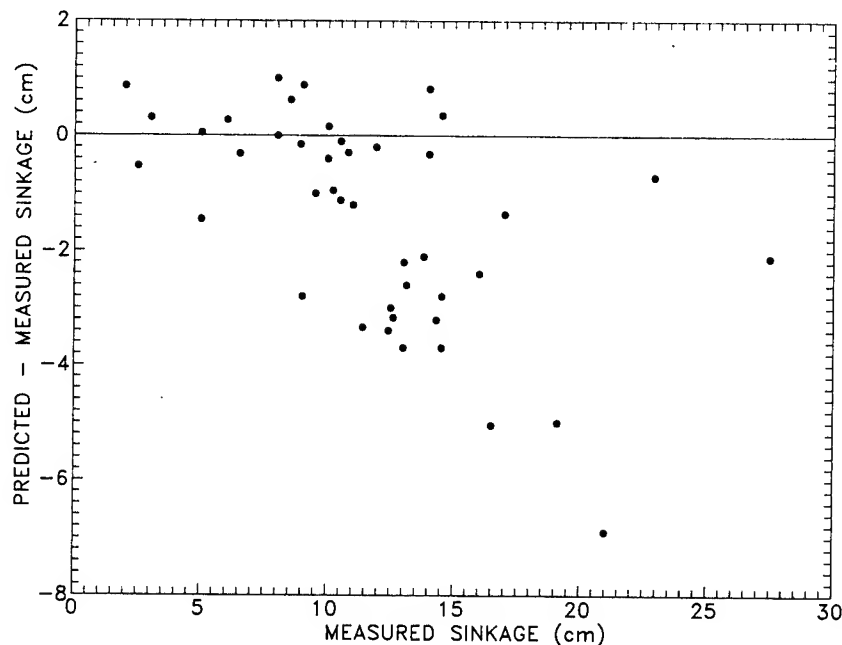


Figure 1. Comparison of measured and predicted sinkage.

Table 4. Vehicle sinkage data 1989.

		Measured density		Snow		Sinkage		
		Width	(kg/m ³)		depth measured		eq 2	
Date		(cm)	Initial	Final	(cm)	(cm)	(cm)	
HMMWV (151.7 kPa)								
$\rho_f = 500$	13 Jan	32	200	380F ^a	15	8	9.0	
			200	480R	15	8	9.0	
			200	470F	19.5	11.9	11.7	
			200	500R	19.5	13.8	11.7	
			200	535F	23.5	19.1	14.1	
			200	—R	23.5	21	14.1	
			200	455F	19.5	14.5	11.7	
			200	470R	17.5	13.1	10.5	
			200	445F	18	14.5	10.8	
			200	475R	18	13	9.9	
	19 Jan		120	320F	13	9	9.1	
			120	410R	12	8.5	14.8	
	20 Jan		120	—F	19.5	14	14.8	
			120	—R	18	14	13.8	
	23 Jan		190	575R	10	9	6.2	
SUSV (13.2 kPa)								
$\rho_f = 500$	13 Jan	60.96	240	415F	14	—	7.3	
			240	450R	15.5	11.4	8.1	
			250	460F	18.5	10.2	9.3	
			250	450R	21	10.8	10.5	
			250	440F	17.5	8.9	8.8	
			250	430R	17	9.5	8.5	
			20 Jan	120	230F	12.5	12.5	9.5
				120	245R	12.5	12.5	9.5
			23 Jan	410	430R	64	19.5	11.5 ^b
				265	455R	54	27.5	25.4
	24 Jan		310	450R	58.4	22.9	22.2	
			450	480R	35.5	5	3.6	

a. F refers to the front axle or track; R refers to the rear axle or track.

b. This data point was not used on Figure 1, because it appears that this represents deep snow rather than shallow snow. A deep snow condition occurs when the disturbed rut snow does not extend down to the underlying layer.

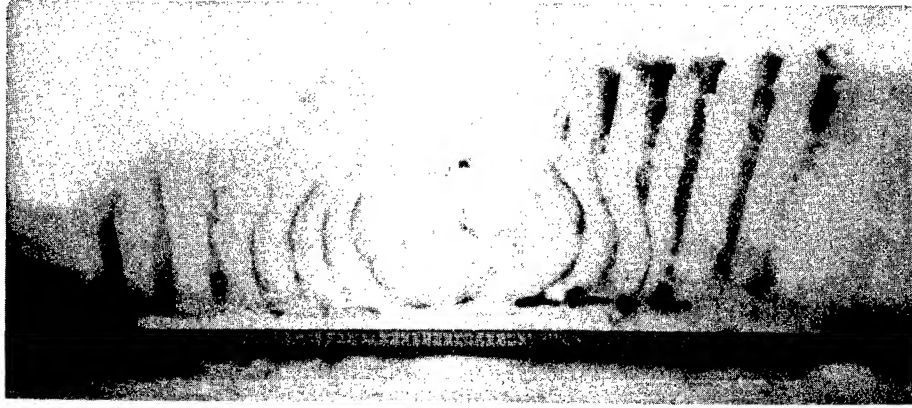


Figure 2. Snow displacement around a vehicle rut.

width of the tire or track, and that the deformed snow is at a uniform density. In an effort to observe how the snow deforms, the snow was marked by filling 1.3-cm (0.5-in.) diameter vertical holes with chalk dust. These holes were aligned perpendicular to the direction of wheel travel, and the vehicle was driven into the marked area. Figure 2 shows a cross section through the chalk dust holes. Clearly, significant lateral snow displacement occurs.

To account for this lateral deformation, the deformed region under the tire rut can be assumed to have an average width (b_a) that is wider than the tire or track width (b). By use of the same volumetric procedure used to develop eq 2, the following is obtained:

$$z = h \left(1 - \frac{b_r}{b_a r_f} \right). \quad (3)$$

Since (b/b_a) is less than one, the effect of this term is to increase the sinkage. This may explain why relatively good predictions of sinkage were obtained with eq 2 when (the consistently high values of) final density from Table 3 were used.

Resistance

As mentioned earlier, motion resistance tests of the CIV were conducted using several different tires to determine the effect of tire width. Comparing the data for tests conducted sequentially in the same test area revealed no clear trend when tire width was the only changing parameter.

The parameter ($\rho_o \times w \times a$) was developed, where a is the length shown in Figure 3, and is estimated using the predicted sinkage (eq 2) and tire or track geometry, w is the maximum tire width and ρ_o is the initial snow density.

The CIV resistance, per tire, is plotted in Figure 4. For the CIV, motion resistance is measured effectively as a one-axle vehicle (trailing tire motion resistance is not measured during a typical CIV motion resistance test). An equation (eq 4, Table 5) of the form $y = dx^e$ fits the data well ($r^2 = 0.76$). Close examination of the data indicates that the data from tire C (the

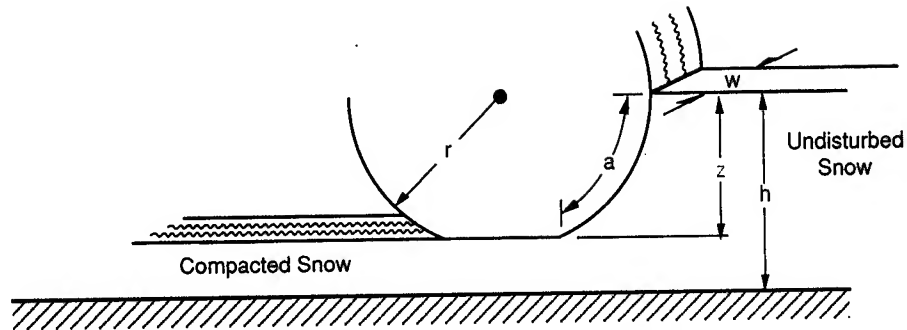


Figure 3. Snow and tire characteristic dimensions.

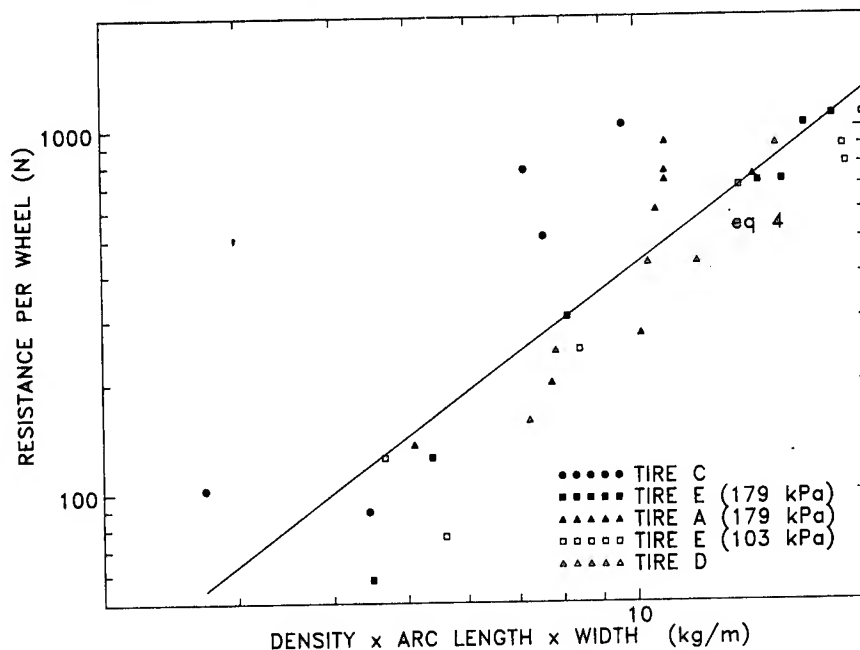


Figure 4. CIV resistance data.

narrowest tire) is near the top of the data range. All the other data, which cannot generally be separated, are for tires that have nearly the same width. This may indicate that the parameter $(\rho_o \times w \times a)$ should be a different function of width, and future work should examine the effect of width more closely. Equation 4 will be used later to compare the CIV data to other vehicles.

Resistance data for several of the wheeled military vehicles (HMMWV, HEMTT, LAV) are shown in Figure 5. Again the parameter $(\rho_o \times w \times a)$ is used. The resistance value used is per tire and thus all vehicle data can be shown on the same graph.

The division of TMR by the total number of wheels or tracks to obtain R_s (from eq 1) is a departure from traditional thoughts on vehicle resistance in snow for multi-axle vehicles. Traditionally, and in the CRREL shallow snow mobility model, it was assumed that trailing

Table 5. Equations obtained from least squares regression analysis.

Equation no	Intercept ^a (d)	Slope (e)	Regression coeff. r^2	Vehicle data used
(4)	11.25	1.58	0.76	CIV
(5)	137.25	0.606	0.44	HMMWV, HEMTT, LAV
(6)	57.63	0.897	0.56	CIV, HMMWV, HEMTT, LAV
(7)	100.79	0.588	0.63	HMMWV
(8)	146.7	0.6	0.67	LAV
(9)	275.43	0.5	0.38	HEMTT
(10)	157.54	0.833	0.27	M113, SUSV, M2

a. The equations are of the form $y = dx^e$, where x is the parameter ($\rho_o \times w \times a$), and y is the external motion resistance due to snow on a per tire or track basis. The SUSV was assumed to have two tracks because of its extremely low ground pressure, and thus the second pair of tracks do not seem to behave as trailing tires.

tires or tracks with lower contact pressures than the preceding tires experience little or no external resistance. By dividing by the number of wheels, I assume that following tires displace as much snow (on a mass basis) as preceding tires or tracks and that this requires an equal amount of effort. Unfortunately, instrumentation limitations did not allow this theory to be examined during these tests.

Figure 5 contains the least squares regression curve (eq 5, Table 5) for the HMMWV, HEMTT, and LAV resistance data; the CIV curve and data are shown for reference. Combin-

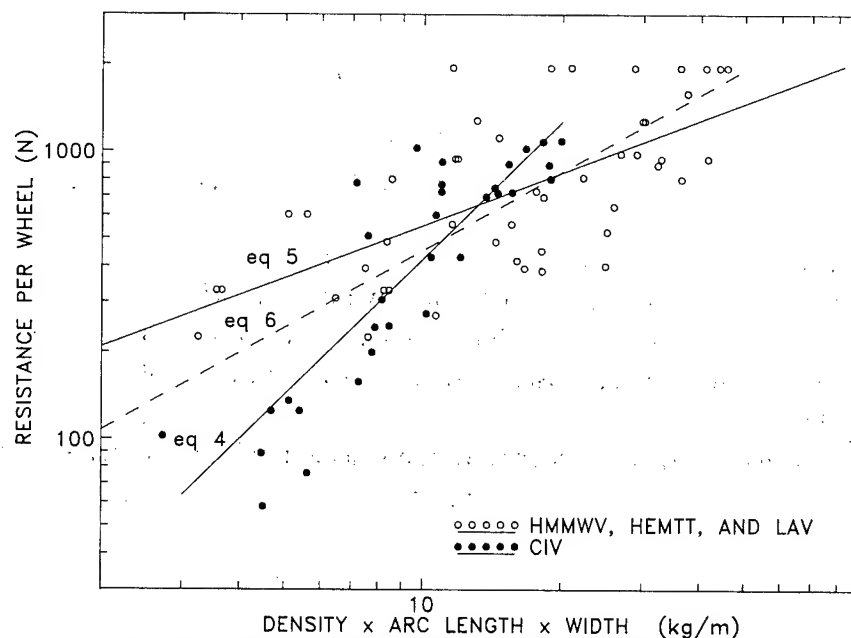


Figure 5. Wheeled vehicle resistance data.

ing the CIV data with the other three wheeled vehicles produced the third regression line on the graph (eq 6, Table 5). Table 5 also contains coefficients for equations obtained for each of the vehicles individually and their corresponding regression coefficients. A trend can be seen in the intercept values of the individual vehicle equations; specifically, they increase with the size of the vehicles. This may be due to increasing tire width or the increase in the number of trailing axles.

Figure 6 compares eq 6 with data obtained from the 5-ton truck during the 1988 field season. Since these data were not used in the development of eq 6, its relative agreement with these "foreign data" is very encouraging.

Figure 7 contains the tracked vehicle resistance data (M113, SUSV, and M2), and again a least squares equation (eq 10, Table 5) is shown. There is much more scatter ($r^2 = 0.27$) in

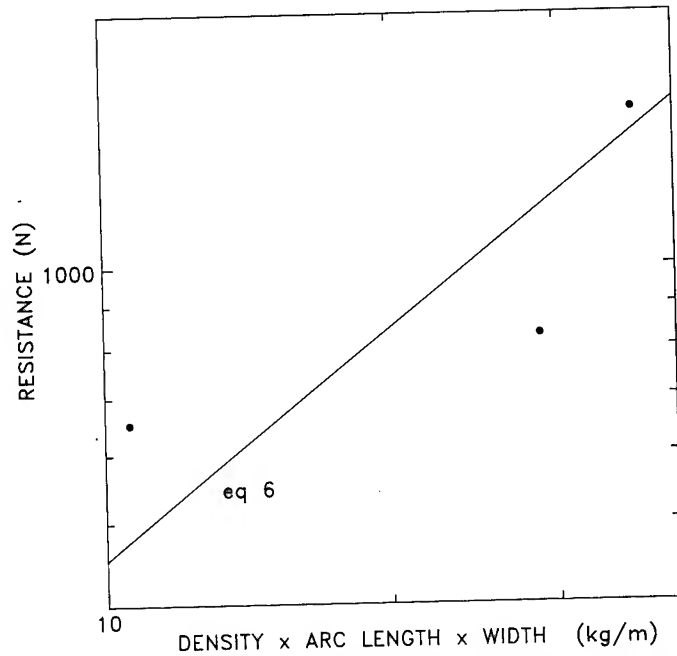


Figure 6. Comparison of eq 6 with 5-ton truck resistance data.

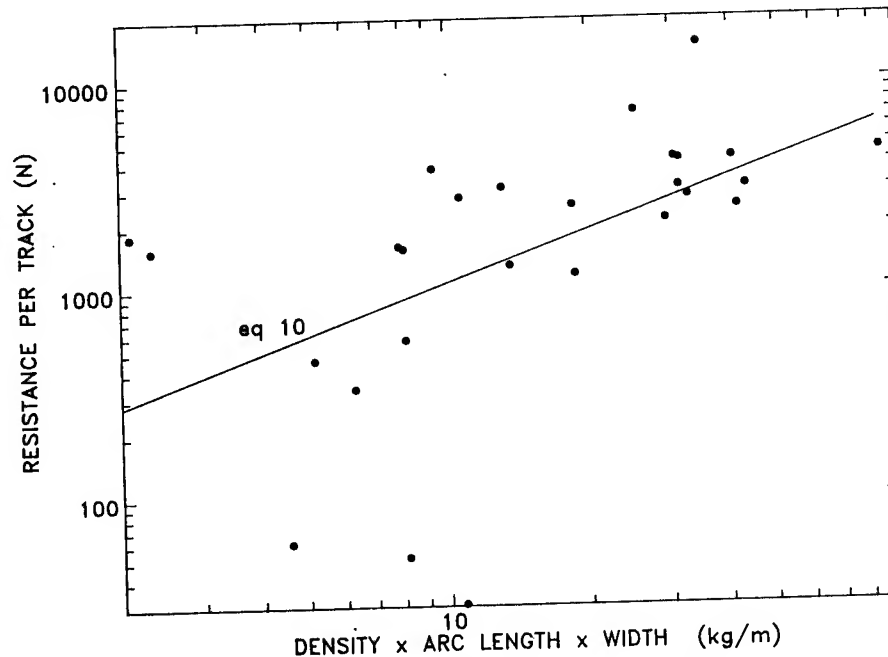


Figure 7. Tracked vehicle resistance data.

these data than in the wheeled vehicle data. The reason for this is not specifically known, further testing of these vehicles is planned in a effort to resolve this question.

Currently, the CRREL shallow snow mobility model uses a method developed by Liston⁵ to predict external motion resistance. A regression curve developed using all the resistance data was compared with the shallow snow mobility model procedure and proved to be no better at predicting motion resistance (average error greater than 50% for both methods.³ Comparisons have not yet been made using the individual equations in Table 5; however, it seems that using one equation for the tracked vehicles and another for the wheeled vehicles should produce more accurate predictions of R_s .

SUMMARY

The sinkage prediction equation used in the CRREL shallow snow mobility model was shown to produce good results.

Empirical resistance equations using a parameter based on simply obtained snow and vehicle parameters were developed that define external vehicle motion resistance. This snow parameter combines what appear to be the important characteristics of the snow/vehicle interaction. Although these equations may not produce a significant improvement in our modeling ability at this time, some trends are becoming evident. Further investigation is required and can be guided by these results.

The assumption used here that trailing tires produce as much resistance as leading tires needs to be confirmed experimentally. The effect of tire width over a greater dimensional range should be examined further in light of the tracked vehicles displaying a much greater resistance than wheeled vehicles and the trends hinted at in the results of the CIV tests with different width tires.

The empirical equations given here should predict external motion resistance well for vehicles with characteristics similar to those tested, as has been demonstrated using the 5-ton truck data.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the many people of CRREL, WES, and KRC who worked long hours in the cold to collect these data, and gives particular thanks to Sally A. Shoop and George L. Blaisdell who reviewed this paper and who are always willing to listen to new ideas.

LITERATURE CITED

1. Blaisdell, G.L., P.W. Richmond, S.A. Shoop, C. Green and R.G. Alger (In prep.) Wheels and Tracks in Snow (Validation Study of the CRREL Shallow Snow Mobility Model). USA Cold Regions Research and Engineering Laboratory, CRREL Report.
2. Green, C.E. and G.L. Blaisdell (In prep.) U.S. Army wheeled versus tracked vehicle snow mobility program. WES/CRREL Technical Report. USA Waterways Experiment Station Special Publication, Vicksburg, Mississippi
3. Richmond, P.W., G.L. Blaisdell, and C.E. Green (In prep.) Vehicle mobility of wheels and tracks in snow (Second validation study of the CRREL shallow snow mobility model). USA Cold Regions Research and Engineering Laboratory, CRREL Report.
4. Blaisdell, G.L. (1983) CRREL instrumented vehicle, hardware and software. USA Cold Regions Research and Engineering Laboratory, Special Report 83-3, 75 pp.
5. Liston, R.A. (1974) Operation of the M151 $\frac{1}{4}$ -ton truck in shallow snow. USA Cold Regions Research and Engineering Laboratory, Technical Note, September (unpublished).

Hyperspectral Imagery - A New Technique For
Targeting and Intelligence

Jack N. Rinker, Dr.
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, VA 22060-5546

Divide and conquer has been the motif of the remote sensing community in its quest for ever more and ever narrower spectral bands to record terrain and target information. Over the years, this has led to an expansion into disparate portions of the electromagnetic spectrum, such as microwave, thermal infrared, photo, and visual, as well as to a contraction of recording bandwidth within any given spectral region. The drive towards narrow bands has been especially successful in the reflected solar radiation portion of the spectrum, i.e., the wavelength range of 0.35-2.5 micra, which represents about 84 percent of the sun's energy.

Why the drive to more and narrower bands? Primarily, by reason of the belief that such will lead to improved detection and identification techniques - a rationale that has much to support it, because the finer the cut of spectral data, the greater the ability to establish identities and conditions.

Because all materials reflect, absorb, or emit photons in ways characteristic of their molecular makeup, a high resolution trace of the intensity of the transmitted, reflected, emitted, or luminesced radiation versus wavelength forms a graphical record unique to a given material. Different materials cannot have identical spectral wave shapes of reflectance, emittance, and luminescence. These characteristic absorption and emission bands occur in narrow wavelength ranges, 10 nanometers (nm) or less; and, unless the instruments have that kind of spectral resolution, these details cannot be recorded. Although many laboratory and field instruments exceed this spectral resolution, airborne systems have only recently entered this domain. Figure 1 shows examples of spectral reflectance measurements and indicates some of the more common absorption bands. From a laboratory point of view, the use of spectral measurements to identify and/or assay components of minerals, pigments, pharmaceutical and other organic and inorganic compounds, is old, established, and reliable. With reference to remote sensing, the reasoning goes that if such could be done from air or space, it would give remote sensing a similar capability.

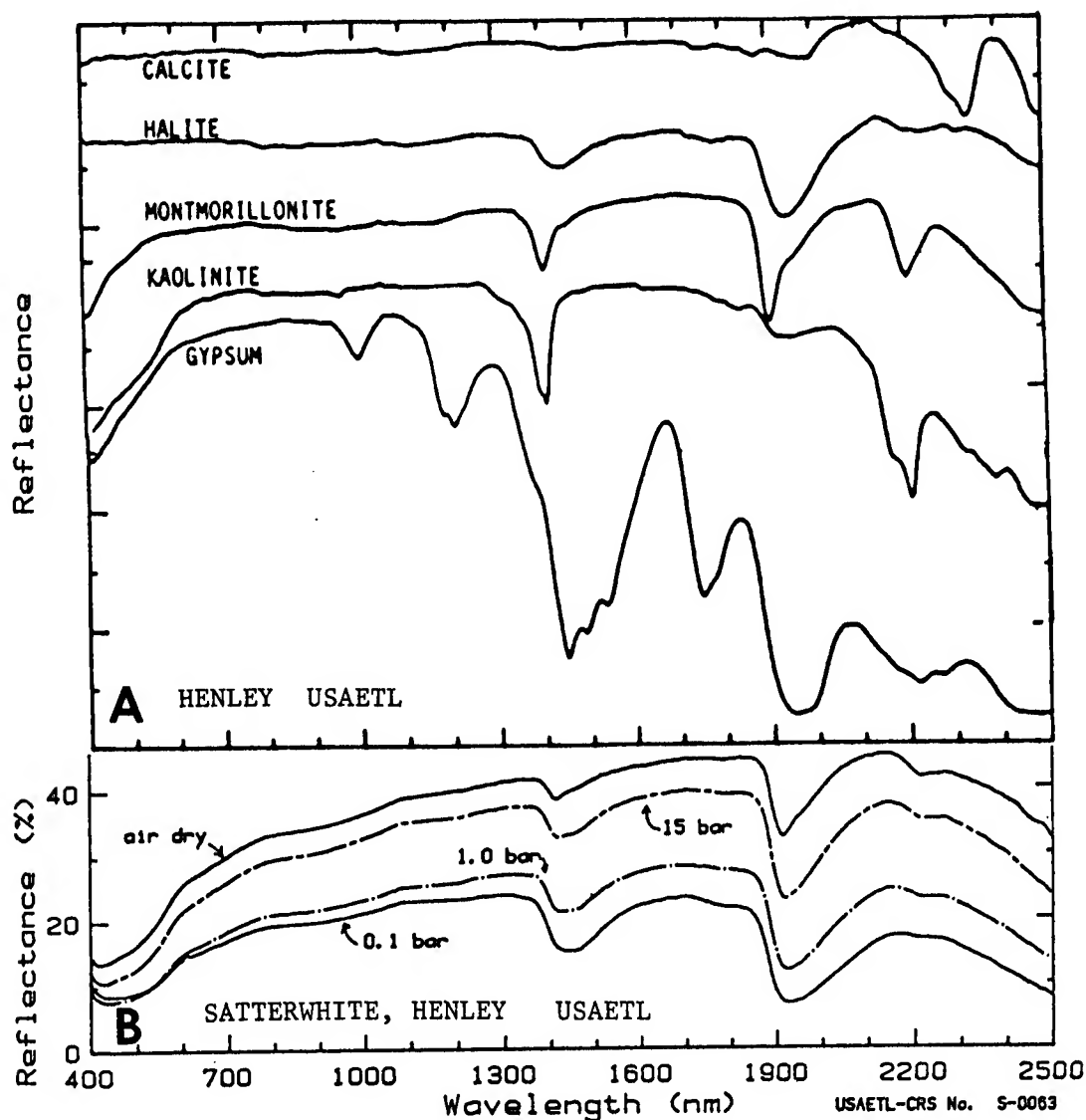


Fig. 1. A. Reflectance measurements of playa surface minerals¹ (records offset vertically to avoid overlap). The strong water absorption bands at 1400 and 1900 nm are apparent. When both are present, it indicates undissociated water, i.e., free water, water of hydration, trapped water in the lattice, etc. Kaolinite shows strong hydroxyl absorption at 1400 and 2200 nm but little at 1900 nm, which suggests a lack of bound water. Molecular water is important in gypsum, and its overtones and combinatorial tones account for the bands at 1000, 1200, and 1700 nm. Calcite is fairly featureless except for the carbonate molecular vibration band at 2300 nm, and sometimes a weak band at 2000 nm. B. Influence of moisture on a silty loam playa soil from Broadwell Lake, San Bernardino, CA, April 1988.²

In relation to systems development, the Army, along with other groups in the 1940s and 50s, divided the photographic portion of the electromagnetic spectrum into narrower bandpasses by using various combinations of photo emulsions and filters³ - i.e., multiband photography. The bandpasses, however, were still broad, ranging from 60 to 100 nm.

Next came the Landsat Multispectral Scanner (MSS), which recorded reflected sunlight in four broad bands, followed by the Thematic Mapper (TM) with six bands in the reflected solar region, and one band in the thermal infrared, with the narrowest band being TM band 3 at 60 nm. Whatever spectral variations occur in the terrain within any of these bands are averaged out to arrive at a digital number (DN) representing the brightness for the whole band. Extensions of the multispectral concept into the thermal infrared region of the spectrum include the Advanced Very High Resolution Radiometer (AVHRR), and the airborne Thermal Infrared Multispectral Scanner (TMS) developed by Daedalus Enterprises, Inc.

In the early 1980s, a system was produced that greatly altered the concepts of multispectral remote sensing with reflected solar energy, i.e., the Airborne Imaging Spectrometer (AIS) developed by the Jet Propulsion Laboratory (JPL).^{4, 5, 6} It recorded reflected solar energy in some 128 channels, or images, within the 1.2-2.4 micra region of the spectrum and with a spectral bandwidth for each channel of less than 10 nm. The AIS evolved into the Airborne Visible Infrared Imaging System (AVIRIS) with some 220 raw data channels, or images, within the 0.4-2.45 micra portion of the spectrum. Resampling gives 210 spectral bands of radiometrically calibrated data. The instantaneous field of view (IFOV) is 1 milliradian, or about 10 meters at operational altitude. Each image is a record of the intensity of reflected sunlight within a spectral bandwidth of less than 10 nm. After calibrations and corrections have been made, the intensity values of the 210 channels, for any given picture element (pixel), can be called up and displayed in sequence along the wavelength axis, as a spectrophotometric trace, i.e., radiometric intensity versus wavelength. Because of the narrowness and multiplicity of the bands, these systems are called hyperspectral, to differentiate them from the broad band systems, e.g., MSS, TM, SPOT, etc. Since then, other narrow bandpass systems have been developed, e.g., the 12 channel imager by Daedalus Enterprises, Inc., and the 64 channel instrument by Geophysical Environmental Research Corp. Planned satellite follow-ons include the Shuttle Imaging Spectrophotometer Experiment (SISEX), and the High Resolution Imaging Spectrometer (HIRIS). Details of these systems can be found in a proceedings issue of the Society of Photo-Optical Instrumentation Engineers.⁷

Figure 2 portrays the hyperspectral concept. The stack of images, 210 in the case of AVIRIS, forms an image cube. The X and Y axes relate to ground, or pixel location, and the third axis to wavelength. Because of the Army's interest in natural and man-made surfaces, it must be able to

work with diverse remote sensor data, and the wavelength axis of its image cube should extend from the ultraviolet, through the reflected solar, thermal infrared, and microwave regions, out to at least L-band radar at about 23.5 cm wavelength, as shown in the figure. Moreover, the Army must be concerned with all photons - reflected, emitted, and luminesced - and be able to move back and forth along the image cube axis, incorporating, evaluating, and comparing whatever imagery bands and other data are available, such as Digital Terrain Elevation Data (DTED). Once the corrected image files are in the computer, the spectral patterns can be evaluated by placing the cursor on the site of interest and bringing up a record of the DN values of each of the involved channels. As a minimum, these data should be able to be displayed as line spectra, intensity versus wavelength spectra, and, in the case of luminescence, as three-dimensional spectra (intensity, excitation, emittance), and as contour plots. The spectra can be evaluated in a number of ways, either directly, or by comparison to a computer library of spectral data bases and models.

Figure 3 is an image cube display of an AVIRIS image over Moffett Field, California. The wavelength axis extends downward. The number 2 on the cube image indicates that it is a record of reflected sunlight in the second, or the 410-420 nm bandpass (approximately). The stack of thin lines parallel to the image and extending downward like pages of a book, are the edges of the other images, i.e., channels 11, 12, 13,... . The strong variations in the edge intensities are due to atmospheric absorption caused mostly by water vapor and oxygen. The dark zone indicated by the arrow consists of four images taken in the atmospheric water absorption bands at 1.35, 1.38, 1.41, and 1.46 micra. With solid and liquid samples, the four bands merge into one broader band centered at about 1.4 micra. See Figure 1. Atmospheric corrections are needed for many targets. If one is interested in vegetation stress, this involves the depths and shapes of a number of water absorption bands. Because water vapor is a component of the atmosphere, the analyst does not know how much of the depth and shape of those water bands is due to atmospheric absorption, and how much is due to vegetation absorption. If corrections can be made to remove the atmospheric component via available models such as LowTran, then the residuum can be attributed to plant water.

An important benefit of an imaging spectrometer is that it provides image patterns and spectral patterns. For terrain information in terms of material identities and conditions, potential for dust generation, engineering site selection and evaluation, probable locations of ground water, subsurface waste disposal, etc., the manual analysis of stereo imagery is still state-of-the-art. For example, an area can be covered with a vegetative mantle, which is all a spectrometer record will show. The stereo shapes of landform and drainage, however, can reveal that beneath the vegetal mantle rests a thinly interbedded series of limestones and shales dipping gently to the west, and with unstable colluvium on the lower slopes.

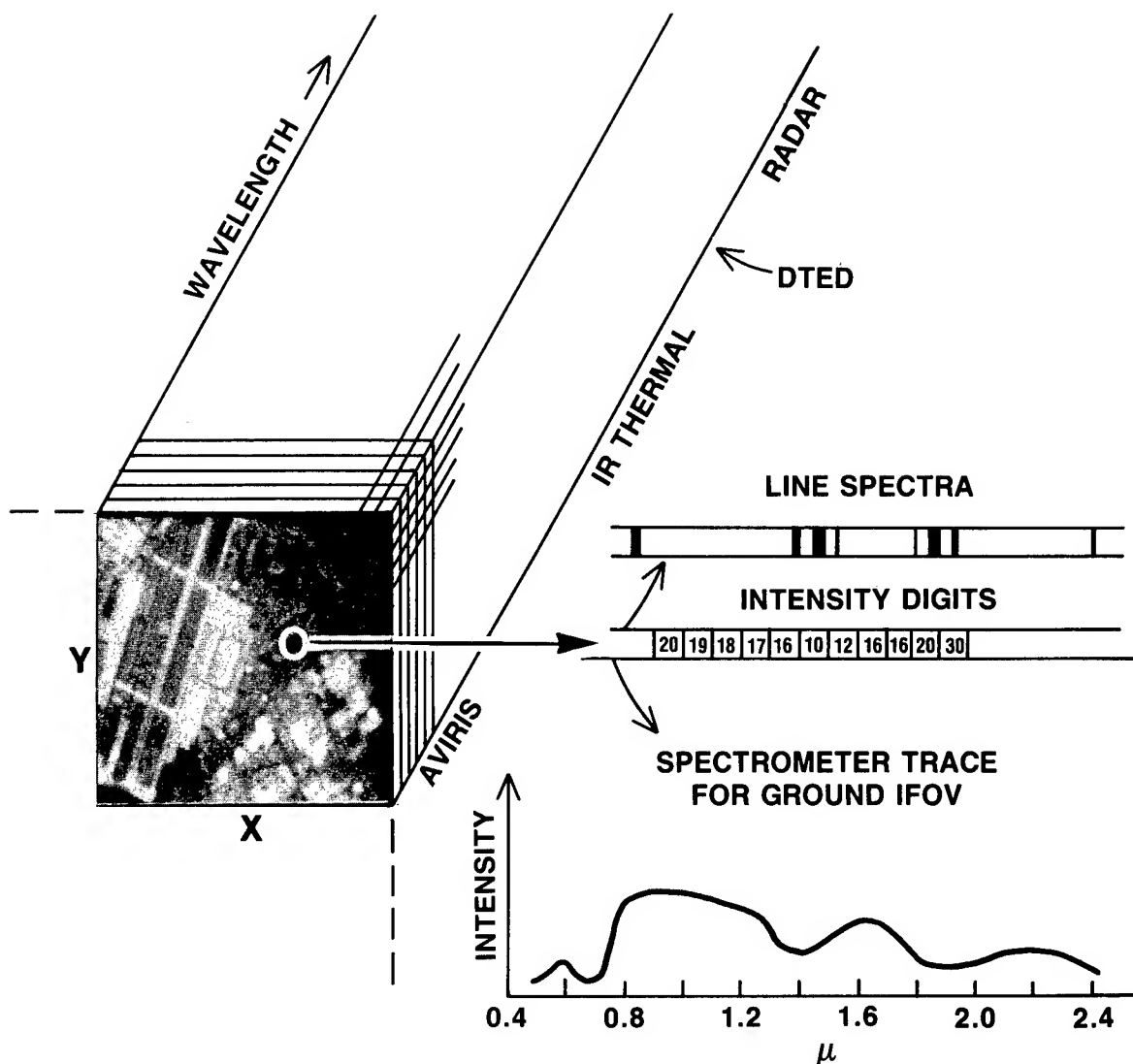


Fig. 2. On the image cube, X and Y indicate ground coordinates, and the third axis is wavelength. For Army purposes it should extend from the ultraviolet out to L-Band radar at 23.5 cm and be able to incorporate other data such as DTED. For any pixel, marked by the cursor, one can sequentially display the intensity values for the bands involved, e.g., MSS, TM, TMS, AVIRIS, HIRIS, etc., as a line display or as a radiance plot of intensity versus wavelength. These can be compared to computer stored spectral data bases to arrive at probable identities. Or, the scene can be searched for all locations that are a spectral match, within some variance range, for a given spectral signature.

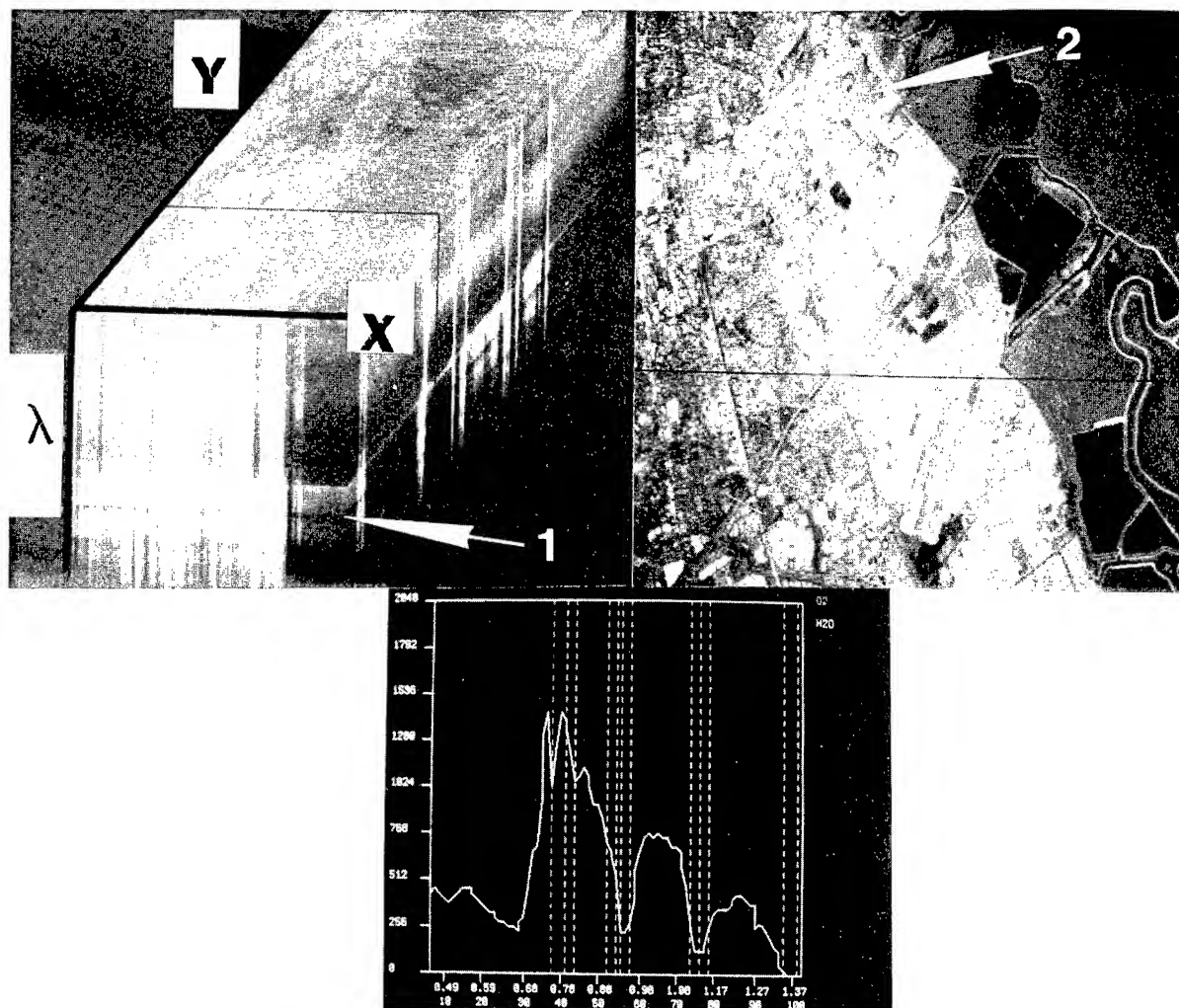


Fig. 3. AVIRIS image of Moffett Field Naval Air Station area, Sunnyvale, CA, taken on 25 June 1987 by JPL. This display was formed on the PIXAR by Barry Holecheck at the U.S. Army Engineer Topographic Laboratories (USAETL). On the image cube the wavelength axis is vertical. The dark zone indicated by arrow no. 1, consists of four images in the water absorption bands at 1.35, 1.38, 1.41, and 1.46 micra. To the right is a black and white print of a color composite image. The graph displays the pixel intensities between 0.4 and 1.37 micra for the vegetated area in the image indicated by arrow no. 2. Although uncorrected for atmospheric absorption, the pattern is typical of chlorophyll. The 12 vertical dashed lines mark atmospheric absorption bands - the left one due to oxygen, and the others due to water vapor. Once corrected, the oxygen band would be eliminated, as well as most traces of the water bands clustered at about .89 and 1.1 micra, and other intensities would be adjusted. The curve would be more similar to plot B at the top of Figure 4.

At present, imaging spectrometers provide only monoscopic imagery, so there is a reduction in the quantity and quality of information that can be derived on the basis of image pattern shapes - but, they are present, and they can make significant direct contributions to an analysis as well as assist in the evaluation of the spectral data. Furthermore, existing routines for combining bands to make color composite images, such as Landsat, Coastal Zone Color Scanner (CZCS), etc., can be applied to AVIRIS imagery as well.

Will this multiplicity of bands lead to data constipation? A good question - and, for the present, the answer is yes, if the data are not properly used. Can the problem be helped by eliminating unneeded bands? Yes, but which ones are unneeded? For targeting minerals, the geologist needs some 30 to 40 bands. For determining crop characteristics, the agriculturalists can get by with perhaps 30 bands, only some of which overlap the geologic needs. The Army, however, with its involvement in terrain use, targeting, cover and concealment, and intelligence, needs information about identities, conditions, and properties associated with vegetation, soils, rocks, minerals, and cultural objects, including camouflage. Doubtless, reductions can be made - but, it is too early for recommendations.

Imaging spectrometer data can be evaluated on the basis of shape of the overall curve, or portions of it, intensity differences/ratios at selected wavelength ranges, wavelength locations of absorption bands, and depths and shapes of absorption bands. To link these to identities requires an extensive library of field and laboratory measurements of spectral reflectance, luminescence, and emittance. Such a library needs excellent documentation, because these measured values change with a variety of factors. For any given surface, the molecular makeup determines the basic characteristics of absorption, reflectance, luminescence, and emittance; but, these are modified by age, weathering, surface structure, orientation, time of day, climate, season, and meteorological variations. For example, vegetation can have smooth, crenulated, or wrinkled leaves of different sizes and arrangements. This means different highlight/shadow ratios, different amounts of transmitted and re-reflected infrared energy through the biomass, and different amounts of radiation reflecting up through the vegetation from the soil surface.⁸ For thermal imagery, changes in incoming short and long wave radiation from space, wind, and atmospheric pressure greatly alter signatures and target/background contrasts. Multiplicity of measurements is necessary because there can be significant variation within any given class of targets, especially in field measurements. For current systems and typical target areas, the IFOV (10 meters for AVIRIS) encompasses a mixture of surfaces, and the resulting spectral signature is a composite of individual signatures - i.e., a mixed pixel, which presents another problem in relation to digital analysis of spectral data.

In any event, the airborne imaging spectrophotometric systems are here, and the concepts are sound. The questions are - what are they suited for? - and, how well will they work?

The first example shows why both reflectance and luminescence spectra are needed. Figure 4-A shows the reflectance characteristics of two fabrics, A and C, and of a typical green leaf, curve B. Fabric C is a reasonable match in the visible and out to about 1100 nm. But, it distorts the 1400 nm water absorption band, shows little mimicry of the 1900 nm band, and would be easy to detect, even with TM imagery if it was large enough. Fabric A mimics the vegetation throughout, including the water absorption bands at 1400 and 1900 nm. It would not be detectable against a vegetal background, and would, in fact, be classified as vegetation. Not detectable! Now what? Well, luminescence techniques for one. Although somewhat neglected at present, such have had a number of successful applications, particularly in the form of the airborne Fraunhofer Line Discriminator (FLD).^{9,10} Figure 4-B shows the luminescence characteristics of these materials. For these measurements, the surface is illuminated with a narrow band of energy at a given wavelength (called "excitation"), and the surface scanned for the spectral distribution of any luminesced photons. This step is repeated at successive wavelength increments of excitation energy until the spectrum of interest has been covered. The result is a three-dimensional plot of excitation wavelength versus emission wavelength versus intensity, similar to Figure 5-B. In Figure 4-B, the luminescence intensities of the fabrics, vegetation, and soil were plotted for the indicated Fraunhofer lines. Fabrics A and C not only have strong signals as compared to soils and vegetation, but the distributions are different. Thus, they would be easily detectable in a soil/vegetation milieu, as well as distinguishable from each other. When bruised, the vegetation showed a strong luminescence that persisted for hours - indicating a possibility for detecting passage of traffic. When the vegetation is pulverized, the luminescence intensity is further increased.

Figure 5-A, shows the reflectance characteristics of four differently dyed areas of a fabric. Although their composite signal would be distinguishable against a background of vegetation, the contrast would not be necessarily strong, particularly against a mixture of soil and different vegetation types and conditions. On a luminescence basis, the fabric has a signal that greatly exceeds that of vegetation, and which occurs at different wavelengths - making detection a certainty if the areal extent is sufficient (graphs B and C).¹¹ The signal threshold for airborne detection is 1,500 units, and the fabric's luminescence peak is 81,000 units. In general, vegetal luminescence rarely reaches 20,000 (senesced conditions), and in the healthy state seldom goes above 12,000. The vegetal sample, corn in this case (graph B), has a peak of 11,000. Soils measured to date show very little luminescence. The iso-intensity contour plot, graph C, shows that the luminescence occurs in different wavelength bands.

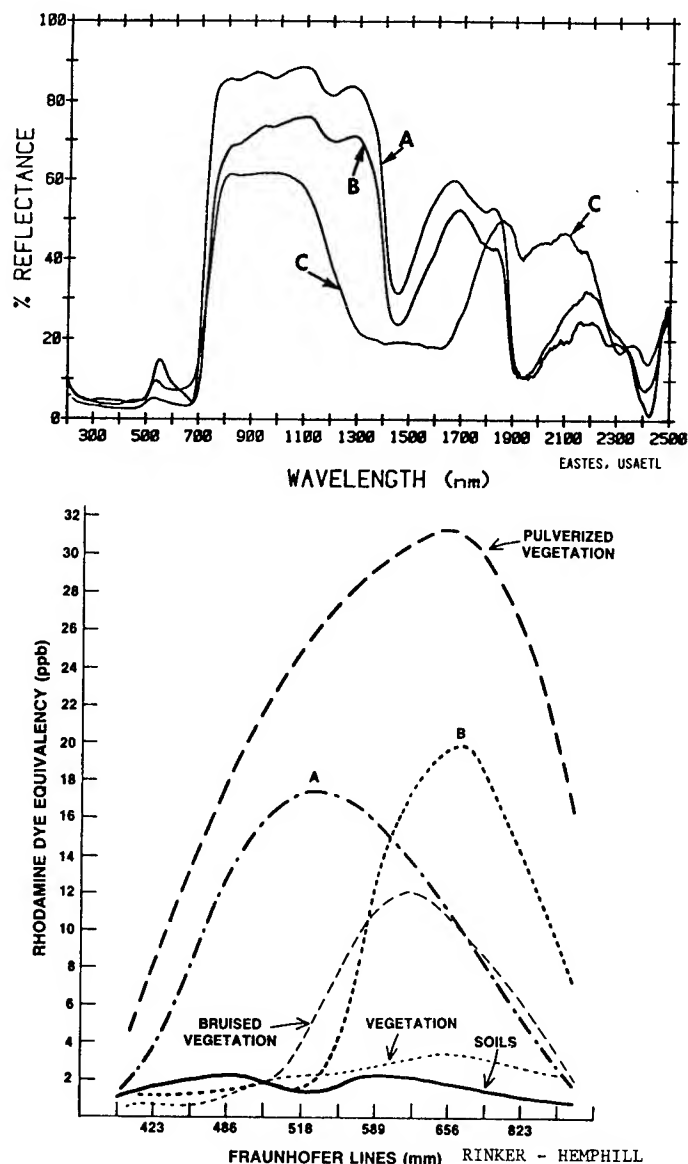


Fig. 4. At the top are spectral reflectances of two fabrics (A and C) and a green leaf (B).¹² Fabric A mimics vegetation, including the water absorption bands, and would be indistinguishable from it. Fabric C is a poor match, and would be easily detected. The lower graph is a plot of luminescence intensity in the Fraunhofer lines of the airborne Fraunhofer Line Discriminator (FLD). Both fabrics show in strong contrast to soils and vegetation (bottom two traces), and also differ significantly from each other. BV indicates the strong signal from bruised vegetation - a signal that persisted for many hours. PV represents the signal from pulverized vegetation.

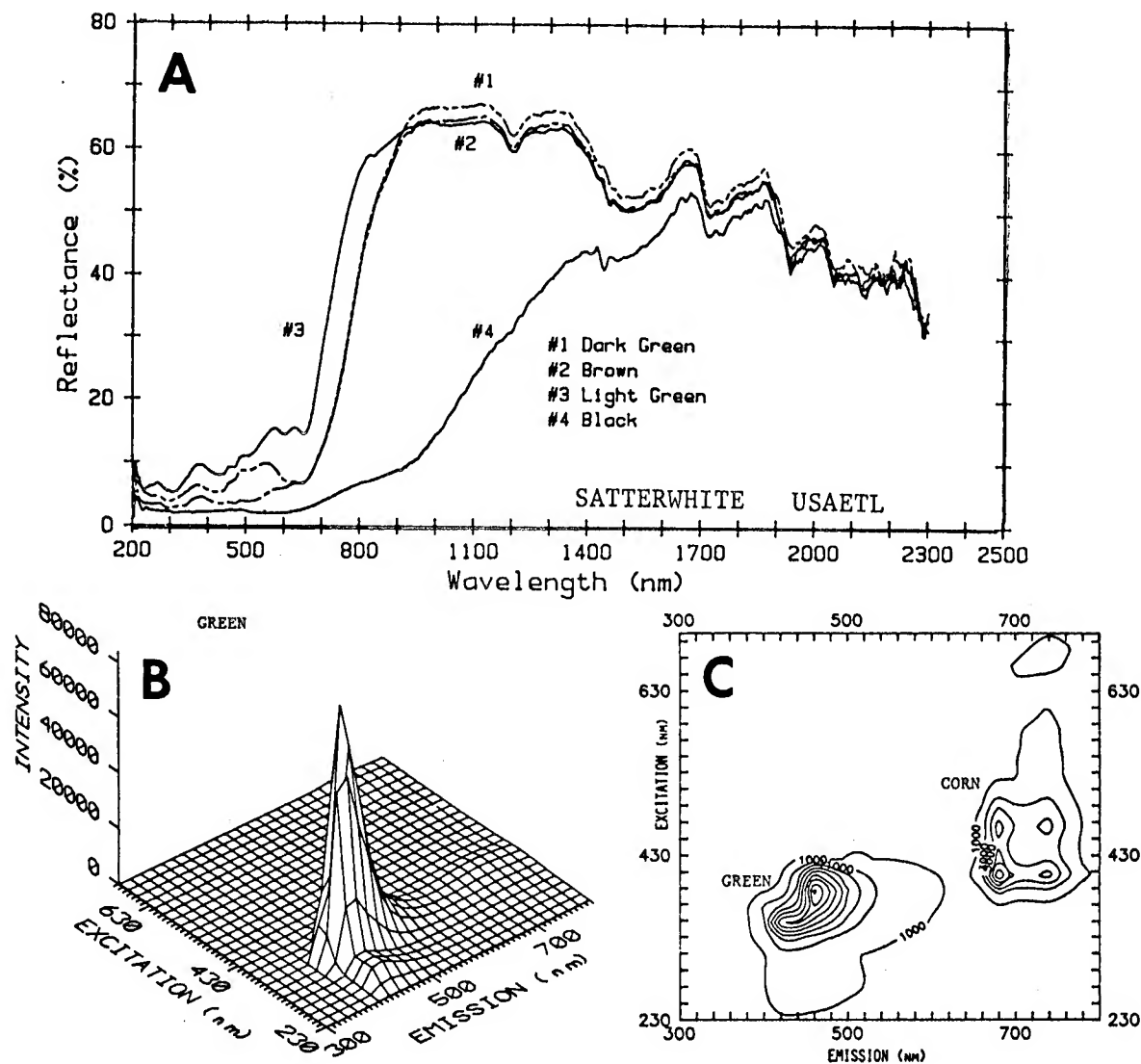


Fig. 5. At the top are the reflectance characteristics of four fabrics. Although distinguishable from vegetation, there would not necessarily be a strong contrast, especially if the background provided a mixture of soils and different types and conditions of vegetation. Graph B is a luminescence plot of the green fabric. Graph C is an iso-intensity plot of the green fabric and of corn, which gives a response typical of healthy herbaceous vegetation. The luminescent intensity of the fabric is 81,000 units. Against the 11,000 intensity units for typical herbaceous vegetation, and the flat response of soils, the fabric with its 81,000 intensity units is easily discernible. Private communication (23 March 1990) suggests a possibility that this high peak is caused by laundry products.

Figure 6 depicts the luminescent characteristics of loblolly pine pollen, whose peak intensities exceed 2,000 units.¹³ Implicit in this illustration are a potential application and a potential problem. First, a possible technique for detecting and monitoring airborne pollen loads, i.e., atmospheric quality. Second is a resulting problem - i.e., such an airborne load can reduce contrast, or otherwise interfere with the recording of terrain surface signals.

Although all materials have spectral reflectance characteristics, they do not all have useful luminescence characteristics. Of that portion we have examined, some general statements can be made. With one exception, soils measured to date do not show useful luminescence. About 75 percent of the vegetal samples and 30 percent of the fabrics have detectable and diagnostic luminescence peaks. For healthy turgid vegetation, these peaks fall in the wavelength range between 640 and 800 nm. As vegetation dries out, these peaks decrease in intensity and peaks develop in the wavelength region between 400 and 600 nm. Intensity distributions are related to material type and condition, and the peak intensities can be sorted into fairly distinct groups based on emission wavelengths. As shown in Figure 7, healthy herbaceous vegetation falls into one assemblage, and everything else, e.g., paints, fabrics, pollen, dry vegetation, senesced vegetation, etc., falls in another. Finer distinctions can be made within these groups.

In support of the hyperspectral program, USAETL is developing classification and analytical software; collecting extensive field and laboratory spectral measurements of soils, rocks, vegetation, and man-made materials; and, in cooperation with the U.S. Geological Survey (USGS) group at Flagstaff, AZ, maintaining a series of instrumented test sites that collect around-the-clock measurements of target/background radiation characteristics and concurrent meteorological conditions.¹⁴ The resulting data bases (Table I) support empirical modeling, assisted target recognition, and digital analysis techniques for hyperspectral imagery.

The data bases and the imagery are but the beginnings. Software is being evaluated and/or developed for incorporating the data bases into the computer library; for sample/target classification hierarchies; for display formats, e.g., line spectra, intensity versus wavelength, three-dimensional plots, contour plots, etc.; and for targeting and analytical procedures. There are basic issues to resolve and tasks to be done that involve all the participants in the interagency hyperspectral working group. These include: establishing variances within sets; determination of significant spectral bands (statistical, mathematical, empirical); the importance of absorption band slope changes; atmospheric backout in relation to targets, areas, and conditions; self-calibration of imagery from known data base sets; mixed pixel problem; influence of steady state illumination on luminescence; and testing and validation of existing models.

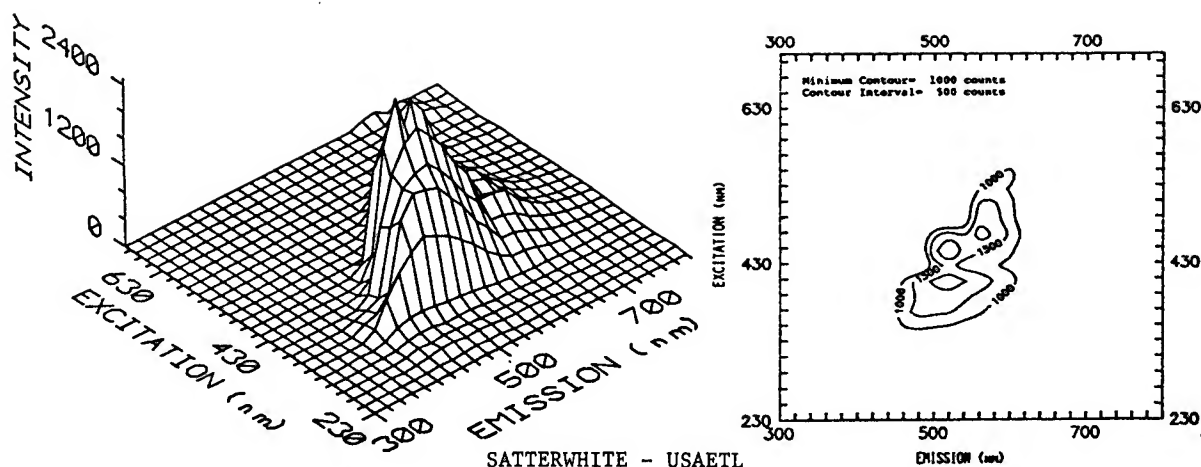


Fig. 6. Luminescence characteristics of loblolly pine pollen. Other pollens also luminesce. Pollen of scrub pine is barely detectable, whereas that of cattails has a signal stronger than that of loblolly pollen. Loblolly pollen, however, consists of very small particle sizes and is readily airborne.

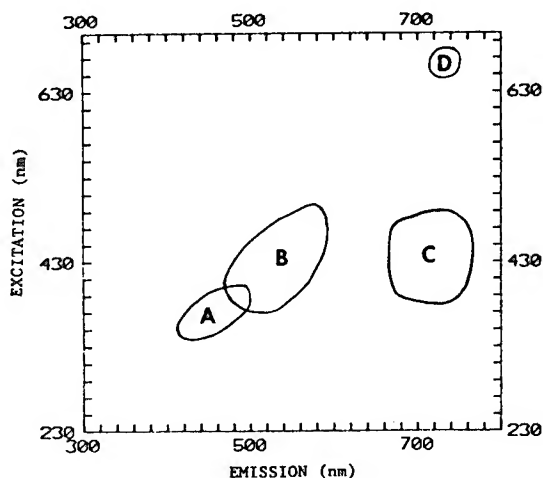


Fig. 7. On a generalized basis, and for those materials that had usable luminescent peaks, the distributions of the peak intensities fall into the indicated areas. With one exception, so far at least, area A contains the fabrics. Area B, which overlaps A to some extent, contains the peaks of pollen, dry vegetation (pine and herbaceous), and senesced vegetation (pine and herbaceous). Area C contains the peaks of healthy vegetation (pine and herbaceous). Area D has a few secondary peaks associated with herbaceous vegetation. The most useful excitation bands are between 330 and 510 nm. The diagnostic emission peaks fall into two groups, 400-600 nm and 660-760 nm, with the latter containing all the healthy herbaceous samples.

Table I. Existing data bases in the USAETL inventory. The temperate data base has eight years of continuous measurements; the subhumid, three and a half years; and the arid, one and a half years. The spectral data bases contain field and laboratory measurements of natural and man-made surfaces, and are collected with an equal or finer spectral resolution than that of the remote sensing hyperspectral systems. Thus, they can be averaged over any selected bandwidth to provide intensity values as they would be in Landsat MSS, TM, TMS, AVIRIS, etc.

ETL Radiation/Meteorological Data Base - Temperate
ETL/USGS Radiation/Meteorological Data Base - Subhumid
ETL/USGS Radiation/Meteorological Data Base - Arid
ETL Spectral Reflectance Data Base - Solar Radiation (0.4-2.5 micra)
ETL Spectral Reflectance Data Base - Thermal Infrared (2.5-14 micra)
ETL Spectral Luminescence Data Base

The airborne systems are here, e.g., AVIRIS, FLD, and TMS, needed spectral data bases exist, and enough hyperspectral image sets (solar reflectance, luminescence, thermal IR) have been evaluated by various interest groups to come to some general conclusions. The National Aeronautics and Space Administration (NASA) and JPL have shown the applicability to targeting minerals. USGS, NASA, and the U.S. Department of Agriculture have shown applicability to minerals, petroleum, and vegetation stress. USAETL has shown applicability to military targeting. No technique is a panacea, no system does everything, and the hyperspectral, even in its broadest sense, has its limitations; but, from the standpoint of targeting, the technique has potential beyond any previous remote sensing endeavor.

Targeting refers to the need to detect, and possibly identify, objects and areas of concern to both military and civil needs. Examples include: military hardware and operational units; items such as traffic use, camouflaged positions, and change detection; vegetation typing and vegetation zones altered by aerosols, chemicals, pathogens, and drought; alteration zones or zones of mineralization; areal extent of damage and boundaries associated with flooding, fire, and other natural and man-induced disasters; sediment loads in water; near water surface phytoplankton distribution; and landscape alterations associated with climatic change. Towards these ends a number of interagency cooperative research efforts have been established, ranging from hyperarid and climatic change studies, through camouflage detection and specialized targeting, to disaster evaluation. The latter endeavor, between USAETL and the Canadian National Transportation Safety Board, arose out of the Army's efforts in applying remote sensing techniques to the Gander, Newfoundland airplane crash in December of 1985.¹⁵

Because of the diversity of requirements within the Army, and the varied activities in its laboratories, the Army has perhaps the largest diverse collection of radiation/meteorological, and spectral reflectance and luminescence data bases available, including associated empirical models and analytical techniques. Although these were developed to support military requirements, they can, without alteration, render direct assistance to critical national and worldwide problems such as narcotics, disaster evaluation, and global climatic change - problems that require all the talents and capabilities that can be brought to bear from both the military and civil domains.

REFERENCES

1. Henley, J.P. 1988. Methods of determining playa surface conditions using remote sensing. Technical Papers of the 1988 ACSM-ASPRS Annual Convention, vol. 6, pp. 108-117.
2. Satterwhite, M.B., and J.P. Henley. Hyperspectral Signatures of Vegetation, Minerals, Soils, Rocks, and Cultural Features - I. Laboratory and Field Measurements. Fort Belvoir, VA: U.S. Army Engineer Topographic Laboratories (in preparation).
3. Rinker, J.N. 1975. Some technical aspects of film emulsions in relation to the analysis and interpretation of aerial photographs. In Aerial Reconnaissance for Archaeology, edited by D.R. Wilson. London: The Council for British Archaeology, Research Report No. 12, pp. 32-46.
4. La Baw, C. 1983. Airborne imaging spectrometer: an advanced concept instrument. Proceedings of Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 430-10.
5. Vane, G. and A.F.H. Goetz (eds). 1985. Proceedings of the Airborne Imaging Spectrometer Data Analysis Workshop. Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology, JPL Publication 85-41, 173 pp.
6. Vane, G. (ed.). 1987. Proceedings of the Third Airborne Imaging Spectrometer Data Analysis Workshop. Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology, JPL Publication 87-30, 183 pp.
7. Vane, G. (ed). 1987. Imaging Spectroscopy II. Proceedings of Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 834.
8. Satterwhite, M.B. and J.N. Rinker. 1986. Time variant reflectance spectra of plant canopies as affected by in-canopy shadows. Proceedings of the Army Science Conference, Vol. IV, pp. 41-55.

9. Hemphill, W.R. and M. Settle (eds). 1981. Workshop on Applications of Luminescence Techniques to Earth Resource Studies. Houston, TX: Lunar and Planetary Institute, LPI Technical Report No. 81-03, 104 pp.
10. Hemphill, W.R., J.N. Rinker, A.F. Theisen, and R.H. Nelson. 1989. Potential Military Applications of Passive Detection of Luminescent Materials (U). Fort Belvoir, VA: U.S. Army Engineer Topographic Laboratories, ETL-RI-1.
11. Satterwhite, M.B. Spectral luminescence of some camouflage materials and paints. 1990 Army Science Conference, June 1990 (in preparation).
12. Measurements provided by John W. Eastes, USAETL.
13. Satterwhite, M.B. Spectral luminescence of plant pollen. IEEE/URSI 10th Annual International Geoscience and Remote Sensing Symposium (IGARSS '90), May 1990 (in preparation).
14. Rice, J.E. and A.E. Krusinger. 1985. A new and extensive thermal contrast data base. Proceedings of IRIS Specialty Group on Targets, Background Discriminations, vol. II, pp. 1-5.
15. Rinker, J.N., J.H. Garstang, D. Edwards, and J. Del Vecchio. 1989. An air photo analysis of an airplane crash. Aviation, Space, and Environmental Medicine, vol. 60, no. 7, supplement, pp. A6-A15.

UNCLASSIFIED

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

Development of Vaccines Against Malaria (U)

Jerald C. Sadoff*, M.D., COL, MC

Daniel Gordon, M.D., MAJ, MC

Anita Aggarwal, Ph.D.

Louis Baron, Ph.D.

Walter Reed Army Institute of Research
Washington, D.C. 20307-5100

Stanley Cryz, Ph.D.

Swiss Serum and Vaccine Institute
Bern, Switzerland

Malaria is the most important infectious disease threat to deployment of US military personnel in major portions of the world. Intensive efforts to eliminate the disease through mosquito control have been failures and hundreds of millions of people are infected.¹ The military has put enormous effort into the development of antibiotics but the parasites have rapidly become resistant. Currently, in large areas of S.E. Asia and Africa there is resistance to almost all but a very few of our drugs. If total antibiotic resistance develops we will not be able to sustain military operations in these areas. Individuals that have lived in non-endemic areas such as the United States have no resistance to malaria and a high morbidity and mortality rate could be expected in the absence of effective treatment. Even under current circumstances where chemoprophylaxis is available infection rates are high in military personnel fighting in malaria endemic areas. During the Viet Nam conflict Malaria was the leading cause of morbidity among military personnel. A vaccine to prevent malaria is therefore one of the highest priorities of the U.S. Army medical research effort.

When a malaria infected mosquito bites a human being malaria sporozoites are regurgitated from the mosquito and rapidly enter the blood stream. They travel to the liver where they

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

invade hepatocytes and macrophages. While in the liver each sporozoite develops into 20,000 merozoites which rupture into the blood stream and invade red blood cells.¹ An attractive way to prevent the disease would be to block the sporozoite from entering the liver cell or to kill it after it has entered the cell but before it becomes 20,000 merozoites. Antibodies against surface components of the sporozoite could block or impair the ability of sporozoites to enter liver cells. Such antibodies could be induced by immunization with sporozoite surface antigens or could be given passively in the form of human monoclonal antibodies. Specific cells known as cytotoxic T cells that are capable of recognizing infected cells and killing the parasites inside them could eliminate sporozoites after they have invaded the liver cells. We have developed vaccines which induce antibodies in humans that block the invasion of sporozoites into liver cells. We have produced a human monoclonal antibody which is also capable of blocking sporozoite invasion of liver cells and could be given for passive protection. We have also designed a novel oral vaccine which induces specific cytotoxic T cells which protect in mouse models against malaria.

The sporozoite has on its surface a molecule known as the circumsporozoite protein. This molecule gives the sporozoite a slimy covering which protects it before it gets into the liver. Mouse monoclonal antibodies against this protein protect mice against malaria infection² and prevent entry of sporozoites into cultured liver cells³. The gene coding for this protein was cloned by Army investigators in collaboration with NIH scientists⁴. The structure of this molecule consists of tandem repeats of four amino acids (NANP) with flanking regions on both amino and carboxy sides of the central repeat. Because the protective monoclonals in mouse malaria systems were directed at the repeat regions of their analogous CS molecules and because humans that were protected from experimental malaria challenge following immunization with irradiated sporozoites appeared to make most of their antibody against the repeat region, the CS repeat was chosen as a vaccine candidate.^{5,6}

The repeat region of the P. falciparum CS protein was cloned into an E. coli expression vector and a recombinant molecule consisting 32 repeats fused to a nonsense protein (tet 32) was produced as a highly purified human vaccine.⁷ This vaccine was highly immunogenic in animals. Unfortunately it was poorly immunogenic in humans. In the 15 volunteers tested the majority made less than 2 micrograms/ml of specific anti-

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

CS antibody. Six immunized volunteers and 2 non-immunized controls all were challenged by allowing five infected mosquitoes to feed on them. The prepatent period or time between mosquito biting and clinical infection in the volunteers was 9 days for one of the controls and 10 days for the other control and the three immunized volunteers with less than 2 micrograms/ml of specific antibody. The prepatent period was 11 days and 12 days for the two volunteers with antibody levels between 2.5 - 5.0 ug/ml. The length of the prepatent period is thought to be related to the level of infection. A longer prepatent period therefore indicates some degree of protection. One volunteer achieved a level of 10 micrograms/ml of specific antibody. This volunteer was completely protected from malaria when challenged.⁸ It therefore appeared that a level of 10 micrograms of specific antibody was required for protection.

A second recombinant vaccine was made by genetically fusing the CS repeat to another protein, the influenza NS1 protein. The goal was to provide better T cell help to enhance the immunogenicity of the repeats. This vaccine although highly immunogenic in animals was also poorly immunogenic in humans. The specific antibody levels were around 2 micrograms/ml. One volunteer had very high levels of antibody with >60 micrograms/ml. This volunteer was challenged with infected mosquitoes and was not protected. Antibody from this volunteer was unable to inhibit sporozoites from invading liver cells in culture (ISI - inhibition of sporozoite invasion). It therefore appeared that vaccines needed to be designed that could induce high enough levels of the right kind of antibody in humans.

We therefore initiated a series of experiments to construct such vaccines by covalent coupling techniques. The variables we manipulated were the carrier protein, the covalent coupling method, the nature of the peptide being coupled, and the ratio of peptide to carrier protein. Methods were developed to couple through the amino or the carboxy terminal or to produce looped peptides. Meningococcal outer membrane proteins, Diphtheria CRM protein, cholera toxin, and Pseudomonas toxin A were covalently coupled to synthetic peptides of the NANP repeat (3 or 6) repeats. Cholera toxin and Pseudomonas toxin A were covalently coupled to a recombinant CS peptide R32LA consisting of 32 repeats with the formula MDP(NANP)₁₅NVDP(NANP)₁₅NVDPLR (M = methionine, D = aspartic acid, P = proline, N = asparagine, A = alanine, V = valine L = leucine, A = arginine).

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

The methods we used for coupling were as follows. The basic reagent we utilized was water soluble carbodiimide, 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDEC) that couples free carboxyl groups to free amino groups. To couple via the carboxyl group of the peptide, carrier proteins were first reacted with a 50 molar excess of adipic dihydrazide (ADH) in 5 molar EDEC at pH 4.8 in water and the pH was kept constant with continuous mixing for three hours. The dihydrazide compound has now reacted with all of the carrier proteins free carboxyl groups. The free peptide is then reacted with the ADH substituted carrier protein. The carboxy terminal region of the peptide reacts with the hydrazine group on the protein in the presence of EDEC about 1000 times more efficiently than with the free amino groups of the peptide so cross linking and polymerization are minimal. In the case of short synthetic peptides molar ratios of peptide to carrier protein in the reaction mixture were 40:1. The reaction was followed by dialysis. In the case of the recombinant R32LA molar ratios were 8:1. The reaction was followed by purification over a sterile Sephadex G75 column and filtration through a 0.45 millipore filter.

To couple via the amino terminal of the peptide the carrier protein was first treated with excess succinic anhydride at pH8 for 30 minutes followed by extensive dialysis. This treatment converts all free amino groups in the carrier protein to carboxyl groups. Peptides are then coupled to the carrier protein in the presence of EDC which couples the free amino terminal group of the peptide to the carrier protein carboxyl group.

The recombinant peptide R32LA was treated with succinic anhydride to convert its free amino terminus to a carboxyl. It was then reacted with ADH treated toxin A. The R32LA was free to couple either by its carboxy or amino terminal region or both.

After appropriate safety testing, and human use approval, vaccines were given to volunteers at a dose of 100 micrograms of carrier protein. The first series of human experiments revealed that independent of carrier protein or coupling method more than 15ug of peptide/100ug of carrier protein was required for human immunogenicity. Vaccines R6MGB(S), R6CTB(S), R6CTB(A), R6CRM(A) all had low peptide:carrier coupling ratios. R6 = a 6 NANP repeat peptide, MGB = Group B meningococcal protein, CTB = Cholera toxin B fragment, CRM = Diphtheria toxin CRM mutant protein (S) = succinylated (A) = treated with Adipic

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

dihydrazide. Immunization with these conjugates produced <0.5 mcg specific antibody/ml.

We were able to make conjugates with higher coupling ratios using *Pseudomonas* toxin A, TA, and cholera toxin, CT, as carrier proteins. Table 1 shows results of immunization of humans at 0 and 8 weeks with these conjugates.

Table 1. Immunogenicity of malaria conjugate vaccines in humans

Vaccine	ratio	ug/100ug	IFA>100	Mean	S.D.
R32TA(A)	7:1	131	3/5	6.0	3.8
R32CT(A)	6:1	95	0/5	2.0	1.3
R32(S)TA(A)	1:1	27.5	1/5	4.9	2.5
R32(S)CT(A)	3:1	43	0/5	3.2	2.4
R6CT(A)	69:1	295	0/5	2.7	0.8
R3CT(A)	25:1	41	1/5	6.2	5.5
R3TA(A)	7:1	14	0/5	1.21	1.1

R32TA(A) which is a covalent conjugate between the recombinant protein R32LA and adipic dihydrazide treated *Pseudomonas* toxin A gave the best overall results. Although R3CT(A) appeared almost as effective as R32TA(A) the immunofluorescent staining of parasites (IFA) and the inhibition of sporozoite invasion of liver cells (ISI) was not as strong. Interestingly when the R32 is succinylated so that it couples by both ends and possibly in a loop 27 ug of peptide was almost as effective as the 131 given with R32TA(A). Although toxin A is a good carrier protein if too little peptide is coupled the results are poor as shown by the lack of immunogenicity of R3TA(A). Based on the results of this experiment we performed an expanded dose response study in humans with R32TA(A). The results of these studies are shown in Table 2. This vaccine was well tolerated even at the highest dose level. No adjuvants other than the widely used alum were required.

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

Table 2. Anti circumsporozoite repeat IgG antibody after immunization of humans with R32TA(A) malaria vaccine.

Dose	day 0	day 56	day 240	IFA > 200
50	0.13 (0.07)	8.1 (5.1)	8.7 (5.0)	1/6
100	0.28 (0.41)	5.6 (7.8)	13.5 (23)	0/5
200	0.08 (0.04)	7.0 (4.0)	9.4 (8.0)	3/6
400	0.10 (0.12)	15.4 (15.8)	26.8 (19.8)	7/7

Mean antibody levels against CSP (standard deviation).

Immunization at day 0, 42 and 226. 25 individuals in the 400 ug group only 7 boosted with the third dose.

The results in Table 2 show that a 400 ug dose of R32TA(A) vaccine induces levels of antibodies in most humans after the second shot which are higher than the 10 ug thought needed for protection. After the third dose mean levels were 2.5 times the protective level. Two volunteers had levels around 8 ug/ml and the other 5 were 25 or greater. These antibodies had strong ability to inhibit sporozoite invasion of liver cells in vitro. The immunized volunteers are currently being challenged with bites from infected mosquitoes. Because of the very promising nature of these results phase II field trials in Kenya and Thailand have been approved and are about to begin.

If military personnel need to deploy rapidly and have not had time to be immunized an alternative is to provide passive immunization by the direct injection of antibody. To solve this problem and also to give us a reagent to guide active vaccine development we produced a human monoclonal antibody from the peripheral blood cells of the one Army volunteer that was protected from malaria following immunization. We reimmunized this volunteer with FSV-1. We collected peripheral blood from him every third day beginning on day 7 after his immunization for 42 days. The cells were EB transformed, allowed to sit in culture for 10-14 days and screened for production of antibody against circumsporozoite protein. Cultures secreting antibody against CS were expanded and then fused to the Heteromyeloma line SHMD-33.⁹ We were able to find peripheral blood cells secreting anti-CS antibody as late as day 35 after immunization which is an unexpected finding. We were only able to stabilize clones that were producing IgM. We

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

cloned a high producer cell line, DGIgM, 7 times. This line has remained stable in continuous culture for over one year. We adapted the cell line to growth in serum free media which makes its production and purification far easier. This antibody is a very potent inhibitor of sporozoite invasion of the liver. The IFA activity was also very high on a microgram basis indicating that this antibody recognizes native sporozoite structures.

The antibody was subjected to Guysen pin analysis. Overlapping peptides from the CS sequence were synthesized on the ends of plastic pins. These pins are then used in an enzyme linked immunoabsorbant assay system to map which peptides the monoclonal antibody reacted best with. Our human monoclonal antibody mapped to PNANPN very strongly but it demonstrated a much broader cross reactivity than an anti CS mouse monoclonal antibody Mab2A10. For example the human monoclonal antibody reacts quite well with ANPNAN while the mouse monoclonal does not. Western blotting studies revealed that DGIgM reacted with sporozoites from P. berghei which is a mouse malaria parasite. Guysen pin mapping revealed that the human monoclonal DGIgM, which was made by immunization with CS peptide from P. falciparum reacted strongly with P. berghei CS sequences. This indicates that important epitopes in the CS molecule may be cross reactive in the two species. To test this we pretreated mice with DGIgM and challenged them with the bites of P. berghei infected mosquitoes. The monoclonal was able to confer about 50% protection against challenge with P. berghei. This suprising finding now provides us with a functional animal test for the potency of this monoclonal.

Human IgM has a half life of around 36 hours while human IgG has a half life of about 26 days. We are therefore having our antibody converted to a human IgG monoclonal by recombinant DNA techniques. The light and heavy variable region domains of DGIgM have been cloned and are being ligated to IgG constant regions. This gene will then be transfected into a eukaryotic myeloma cell line capable of synthesis and assembly of human immunoglobulin. This antibody will then be tested for potency in human challenge studies and if protective will be produced for fielding in the Army. If this monoclonal proves effective it will be a guide to help enhance our development of active vaccines. Such a reagent would be extremely helpful in telling us whether carboxy, amino or looped coupling is optimal. Likewise a synthetic peptide able to exactly mimic DGIgM's binding site may be an even better active vaccine.

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

Under experimental conditions humans have been protected from P. falciparum malaria by immunization with irradiated sporozoites. Mice can also be protected by immunization with irradiated sporozoites. It has been demonstrated that irradiated sporozoites can induce protection in mice that have no B cells and are therefore incapable of producing antibody. Cellular immunity is all that is required for sporozoite induced immunity.¹⁰ We believed that intracellular targeting of malaria antigens, preferably to cells within the liver was required for induction of this cellular immunity. We reasoned that if we put malaria genes into a bacteria that could invade liver cells and produce malaria antigens inside these cells it could fool the immune system to think that an actual malarial intracellular infection had occurred. Salmonella typhimurium is a bacterial organism that after being orally ingested rapidly makes its way from the intestine to the lymphatic system and the liver where it invades cells and survives in the intracellular environment. We have used avirulent S. typhimurium strains such as WR4017 to deliver and produce malaria antigens intracellularly. These bacterial recombinant strains are able to invade cells but do not replicate very well or for very long inside macrophages or other cells and therefore do not cause disease.

Using molecular biology techniques we transformed the gene coding for the full length P. berghei circumsporozoite protein (repeat plus flanking regions) into S. typhimurium WR4017. Using sensitive ELISA techniques we were able to detect malaria CS protein being produced by the bacteria. We then fed these live recombinant bacterial vaccines to mice. Controls consisted of mice orally immunized with Salmonella WR4017 that did not contain malaria genes. Nine weeks later we injected sporozoite lysate into the footpads of these two groups of mice. Mice orally immunized with the Salmonella-P. berghei CS recombinants had significant swelling compared to controls (0.00 ± 0.00 vs 0.12 ± 0.036 $p < 0.01$). This indicated that these vaccines were able to induce delayed type hypersensitivity, an important cellular immune response.¹¹

Five weeks and nine weeks after oral immunization with a single dose of 10^9 Salmonella-P. berghei-CS recombinants 70% protection against challenge with 1500 P. berghei sporozoites was seen. This protection was seen in the absence of antibody and was the first demonstration that the CS protein was a target of cellular immunity.¹¹

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

Because the protection was achieved in the absence of antibody we proposed that cell mediated immune mechanisms were involved. We postulated that intracellular targeting of the *Salmonella* recombinant organisms led to expression of CS antigen on the cell surface in association with class I MHC molecules with the subsequent induction of specific class I restricted CD8⁺ cytotoxic T lymphocytes (CTLs). To test this hypothesis we constructed new *Salmonella typhimurium*-*P. berghei* CS and *S. typhimurium*-*P. falciparum*- CS recombinants, immunized mice and tested for CS specific CTLs. We also performed *in vivo* depletion experiments to determine if CD8⁺T cells mediated the vaccine induced protection.

Five weeks after oral immunization with *Salmonella*-*P. berghei* CS recombinant WR4024/pMGB2 which expresses the full length *P. berghei* CS protein 12 of 22 mice (55%) were protected from lethal challenge. After depletion of CD8⁺ T cells in the mouse by *in vivo* treatment with anti CD8 monoclonal antibody only 1 of 19 mice previously immunized with WR4024/pMGB2 was protected. Depletion of CD8⁺ T cells therefore completely abrogated the immunity induced by the *Salmonella* recombinant vaccine. Mice depleted of CD4⁺ cells did not lose their immunity. This is strong evidence that specific CD8⁺ cytotoxic T cells are induced by these vaccines and are responsible for the protection seen.

To further characterize the CD8⁺ T cells induced by these live oral vaccines *in vitro* cytotoxic T cell assays were performed. We utilized the methods of Townsend et. al. where target cells are made by incubation of histocompatible cells with short synthetic peptides which associate with class I molecules.¹² These cells then look like cells that have intracellular pathogens that have processed antigens and presented them on their surface. We designed peptides from the *P. berghei* CS protein based on the sequence identified previously by Romero et al.¹³ Five different peptides from the amino terminus, repeat region and carboxy terminus were used. Spleen cells were taken 4-10 weeks after immunization. After *in vitro* stimulation with peptide for 6 days killing of peptide labelled target cells was measured by ⁵¹ Chromium release. Spleen cells from mice immunized with WR4024/pMGB2 caused 24% specific lysis of histocompatible target cells incubated peptide 242-253 (NDDSYIPSAEKI). Cells from control mice and unstimulated cells did not cause lysis. Cells from immunized mice did not lyse cells incubated with any other peptide. *In vitro* CD8 depletion experiments proved that the cells causing chromium release were CD8⁺ cells. No specific lysis was seen

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

when peptides were incubated with non-histocompatible cells (EL4(H-2^b) proving that the cytolytic activity was genetically restricted.

The peptide recognized by these class I CD8⁺ CTLs which have been induced by these oral Salmonella recombinant vaccines is the same as the peptide recently identified by Romero as the target of CTLs in mice that have been immunized with irradiated sporozoites. These recombinant vaccines therefore induce some of the same CTL mechanisms of immunity induced by irradiated sporozoites. Remarkably they are directed at the same peptide region in the CS molecule.

Using similar techniques we demonstrated that mice orally immunized with live S. typhimurium-P. falciparum CS recombinant vaccines produced Class I restricted CD8⁺ CTLs directed against epitopes found in the flanking region of CS protein. These CTLs mapped to peptide 291-310 (DELDYENDIEKKICKMEKCSS) which is the same peptide previously identified as the target of CTLs induced by sporozoite or recombinant-Vaccinia immunization. These specific CTLs could be induced by Salmonella vaccines that contained full length CS gene or a repeatless version missing all of the middle repeat region. A fully intact molecule is therefore not required for induction of CD8⁺ CTLs.

Salmonella-malaria recombinant vaccines are therefore capable of inducing cytotoxic mediated immunity and protection against malaria. We are currently constructing vaccines for human use in S. typhi mutants suitable for human use. To stabilize these genes we have adapted methods for chromosomal insertion. These techniques eliminate antibiotic markers as well. Four other genes, liver stage antigen 1, heat shock protein, CSP11 and RISA are being inserted into Salmonella to more closely simulate the variety of antigens presented by the sporozoite and the liver stage parasite. Salmonella typhi Ty21 is the attenuated carrier strain we are currently using for construction of human vaccines. This strain has been used as an oral vaccine to prevent typhoid fever in over 20 million people with no difficulties. The inherent safety of these genetically attenuated bacteria that are susceptible to antibiotics, plus their usefulness as oral agents makes them very attractive as carriers. Our finding that they induce CTL mediated effector mechanisms suggest that live attenuated recombinant Salmonella vaccines may be useful against other military important diseases where CTL mediated immunity may be important such as HIV and bacterial and viral biological threat agents.

UNCLASSIFIED

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

We believe that an effective vaccine against malaria will have an antibody as well as a cellular component. We have developed a vaccine which induces high levels of antibody in humans which may offer significant protection to the soldier and have developed a human monoclonal antibody which may afford immediate protection when appropriate. We have also laid the foundation and discovered a potentially practical, safe, and new class of oral vaccines for induction of cellular immunity against malaria. All of these products will be in challenge or field trials in the next year.

UNCLASSIFIED

SADOFF, GORDON, AGGARWAL, BARON, CRYZ

References:

1. Miller LH, Russell JH, Carter R, Good MF, Nussenzweig V, Nussenzweig RS. Science, 234, 1349, (1986).
2. Potocnjak P, Yoshida N, Nussenzweig RS, Nussenzweig V. J. Exp Med., 151, 1504 (1980).
3. Hollingdale M, Zavala F, Nussenzweig RS, Nussenzweig V. J. Immunol., 128, 1929, (1982).
4. Dame JB, Williams JL, McCutchan TF, Science, 225, 593 (1984).
5. Nussenzweig V, Nussenzweig RS, Adv. Immunol., 45, 283, (1989).
6. Ballou WR, Rothbard J, Wirtz RA et. al. Science, 228, 996, (1985).
7. Young JF, Hockmeyer WT, Gross MT et. al. Science, 228, 958 (1985).
8. Ballou WR, Sherwood JA, Neva FA, et. al., Lancet, 8545, 1277 (1987).
9. Teng,NNH, Kaplan HS, Heber JM, Moor C, Douglas H, Wunderlich A, Braude AI, Proc Natl Acad Sci USA. 82, 1790, (1985).
10. Chen DH, Tigelaar, RE, Weinbaum FI. J. Immunol, 118,1322 (1977)
11. Sadoff JC, Ballou WR, Baron LS, et al., Science, 240, 336, 1988.
12. Townsend ARM, Rothbard FM, Gotch, Bahadur D, Wraith D, McMichael. Cell. 44,959, (1986).
13. Romero P, Maryanski JL, Corradin G, Nussenzweig RS, Nussenzweig V, Zavala F. Nature, 341, 323, (1989).
14. Kumar S, Miller LH, Quakyi DB, Keister RA, Houghten WL, Maloy B, Moss JA, Berzofsky, Good MF, Nature, 334, 258, (1988).

Three Dimensional Flow Calculations for a Projectile
with Standard and Dome Bases

Jubaraj Sahu and Charles J. Nietubicz
US Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

I. INTRODUCTION

The ability to compute the base region flow field for projectile configurations using Navier-Stokes computational techniques has been developed over the past few years^{1,2,3}. This capability is very important for determining aerodynamic coefficient data including the total aerodynamic drag. The majority of base flow calculations to date have modeled the base region as a flat solid surface. Many of the actual configurations have some form of base cavity. General opinion has been that the inclusion of a base cavity or modifications to the interior cavity of a projectile base would have little or no effect on the overall flight performance parameters.

The M825 projectile under certain conditions is expected to be aeroballistically similar to its parent configuration the M483A1. The M825 has an aluminum/steel base which is configured as a flat cavity (standard). A recent Product Improvement Program (PIP), undertaken to reduce the production costs and improve shell integrity, resulted in the design of a new base configuration. This new PIP configuration has an all steel base and contains a dome cavity. A series of aeroballistic tests⁴ were conducted in the Transonic Range Facility of the US Army Ballistic Research Laboratory (BRL), to determine any difference in the aeroballistics which may occur between the standard and dome base configurations. As a result of these tests, differences in aerodynamic performance were found to exist between the two rounds. The most significant changes in the aerodynamic data were in the lift and static moment coefficients. The drag was found to differ by a few percent with the dome configuration having the lower drag at low transonic speeds.

A computational study was undertaken to determine the ability of the present Navier-Stokes codes to predict these differences and to further understand the fluid

dynamic behavior which can account for such small changes. The use of Navier-Stokes codes can provide a detailed description of the flow field associated with the M825 configuration as well as the integrated aerodynamic coefficients. The initial results for the zero degree angle of attack case have been reported by Sahu et. al.⁵. This was accomplished using an axisymmetric base flow code and the results showed the same effect as the range data, that is, a small reduction in the total aerodynamic drag at low transonic speeds ($M < 0.95$) for the dome base configuration. The trend reversed at high transonic speeds ($M > 0.98$). This paper describes an extension of that work into three dimensions.

Numerical computations have been performed using a 3D zonal, implicit Navier-Stokes code. The Mach number range was $0.8 \leq M \leq 1.5$ for an angle of attack, $\alpha = 4.0^\circ$. Results presented include the qualitative features of the base region flow field for the two base cavities. Aerodynamic force and moment coefficients have been obtained from the computed solutions and are presented as a function of Mach number. Computed results show small differences in normal force and pitching moment coefficients similar to that found in the range data.

II. GOVERNING EQUATIONS AND SOLUTION TECHNIQUE

The complete set of time-dependent thin-layer Navier-Stokes equations is solved numerically to obtain a solution to this problem. The numerical technique used is an implicit finite difference scheme. Although time-dependent calculations are made, the transient flow is not of primary interest at the present time. The steady flow, which is the desired result, is obtained in a time asymptotic fashion.

1. GOVERNING EQUATIONS

The complete set of three dimensional, time dependent, generalized geometry, thin-layer, Navier-Stokes equations for general spatial coordinates ξ, η, ζ can be written as⁶:

$$\partial_\tau \hat{q} + \partial_\xi \hat{F} + \partial_\eta \hat{G} + \partial_\zeta \hat{H} = Re^{-1} \partial_\zeta \hat{S} \quad (1)$$

where

$$\begin{aligned} \xi &= \xi(x, y, z, t) && \text{- longitudinal coordinate} \\ \eta &= \eta(x, y, z, t) && \text{- circumferential coordinate} \\ \zeta &= \zeta(x, y, z, t) && \text{- nearly normal coordinate} \\ \tau &= t && \text{- time} \end{aligned}$$

and

$$\begin{aligned}
\hat{q} &= \frac{1}{J} \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ e \end{bmatrix} & \hat{F} &= \frac{1}{J} \begin{bmatrix} \rho U \\ \rho u U + \xi_x p \\ \rho v U + \xi_y p \\ \rho w U + \xi_z p \\ (e + p)U - \xi_t p \end{bmatrix} \\
\hat{G} &= \frac{1}{J} \begin{bmatrix} \rho V \\ \rho u V + \eta_x p \\ \rho v V + \eta_y p \\ \rho w V + \eta_z p \\ (e + p)V - \eta_t p \end{bmatrix} & \hat{H} &= \frac{1}{J} \begin{bmatrix} \rho W \\ \rho u W + \zeta_x p \\ \rho v W + \zeta_y p \\ \rho w W + \zeta_z p \\ (e + p)W - \zeta_t p \end{bmatrix}
\end{aligned} \tag{2}$$

and where

$$\hat{S} = \frac{1}{J} \begin{bmatrix} 0 \\ \mu(\zeta_x^2 + \zeta_y^2 + \zeta_z^2)u_\zeta + \frac{\mu}{3}(\zeta_x u_\zeta + \zeta_y v_\zeta + \zeta_z w_\zeta)\zeta_x \\ \mu(\zeta_x^2 + \zeta_y^2 + \zeta_z^2)v_\zeta + \frac{\mu}{3}(\zeta_x u_\zeta + \zeta_y v_\zeta + \zeta_z w_\zeta)\zeta_y \\ \mu(\zeta_x^2 + \zeta_y^2 + \zeta_z^2)w_\zeta + \frac{\mu}{3}(\zeta_x u_\zeta + \zeta_y v_\zeta + \zeta_z w_\zeta)\zeta_z \\ \{(\zeta_x^2 + \zeta_y^2 + \zeta_z^2)[\frac{\mu}{2}(u^2 + v^2 + w^2)_\zeta \\ + \frac{\kappa a_\zeta^2}{Pr(\gamma - 1)}] \\ + \frac{\mu}{3}(\zeta_x u + \zeta_y v + \zeta_z w)(\zeta_x u_\zeta + \zeta_y v_\zeta + \zeta_z w_\zeta)\} \end{bmatrix} \tag{3}$$

In equation (1), the thin-layer approximation is used and the viscous terms involving velocity gradients in both the longitudinal and circumferential directions are neglected. The viscous terms are retained, however, for velocity gradients in a direction nearly normal to the surface where large flowfield gradients exist. These viscous terms in ζ are collected into the vector \hat{S} .

For this computation, the diffusion coefficients μ and κ contain molecular and turbulent parts. The turbulent contributions are supplied through an algebraic eddy-viscosity hypothesis which has been developed by Baldwin and Lomax.⁷

The velocities in the ξ , η , and ζ coordinate directions can be written

$$\begin{aligned} U &= \xi_t + u\xi_x + v\xi_y + w\xi_z \\ V &= \eta_t + u\eta_x + v\eta_y + w\eta_z \\ W &= \zeta_t + u\zeta_x + v\zeta_y + w\zeta_z \end{aligned}$$

which represent the contravariant velocity components.

The Cartesian velocity components (u , v , w) are retained as the dependent variables and are nondimensionalized with respect to a_∞ (the free stream speed of sound). The local pressure is determined using the relation

$$p = (\gamma - 1)[e - 0.5\rho(u^2 + v^2 + w^2)] \quad (4)$$

where γ is the ratio of specific heats. Density (ρ) is referenced to ρ_∞ and the total energy (e) to $\rho_\infty a_\infty^2$. The transport coefficients are also nondimensionalized with respect to the corresponding free stream variables. Thus the Prandtl number which appears in \hat{S} is defined as $Pr = c_{p\infty}\mu_\infty/\kappa_\infty$.

In differencing these equations it is often advantageous to difference about a known base solution denoted by subscript $_0$ as

$$\begin{aligned} \delta_\tau(\hat{Q} - \hat{Q}_0) + \delta_\xi(\hat{F} - \hat{F}_0) + \delta_\eta(\hat{G} - \hat{G}_0) + \delta_\zeta(\hat{H} - \hat{H}_0) - Re^{-1}\delta_\zeta(\hat{S} - \hat{S}_0) \\ = -\partial_\tau\hat{Q}_0 - \partial_\xi\hat{F}_0 - \partial_\eta\hat{G}_0 - \partial_\zeta\hat{H}_0 + Re^{-1}\partial_\zeta\hat{S}_0 \end{aligned} \quad (5)$$

where δ indicates a general difference operator, and ∂ is the differential operator. If the base state can be properly chosen, the differenced quantities can have smaller and smoother variation and therefore less differencing error.

2. NUMERICAL TECHNIQUE

The implicit approximately factored scheme for the thin layer Navier Stokes equations that uses central differencing in the η and ζ directions and upwinding in ξ is written in the form

$$\begin{aligned} & \left[I + h\delta_\xi^b(\hat{A}^+)^n + h\delta_\zeta\hat{C}^n - hRe^{-1}\bar{\delta}_\zeta J^{-1}\hat{M}^n J - D_i|_\zeta \right] \\ & \quad \times \left[I + h\delta_\xi^f(\hat{A}^-)^n + h\delta_\eta\hat{B}^n - D_i|_\eta \right] \Delta\hat{Q}^n = \\ & -\Delta t \{ \delta_\xi^b[(\hat{F}^+)^n - \hat{F}_\infty^+] + \delta_\xi^f[(\hat{F}^-)^n - \hat{F}_\infty^-] + \delta_\eta(\hat{G}^n - \hat{G}_\infty) \\ & \quad + \delta_\zeta(\hat{H}^n - \hat{H}_\infty) - Re^{-1}\bar{\delta}_\zeta(\hat{S}^n - \hat{S}_\infty) \} - D_e(\hat{Q}^n - \hat{Q}_\infty) \end{aligned} \quad (6)$$

where $h = \Delta t$ or $(\Delta t)/2$ and the free stream base solution is used. Here δ is typically a three point second order accurate central difference operator, $\bar{\delta}$ is a midpoint operator used with the viscous terms, and the operators δ_ξ^b and δ_ξ^f are backward and forward three-point difference operators. The flux \hat{F} has been eigensplit and the matrices \hat{A} , \hat{B} , \hat{C} , and \hat{M} result from local linearization of the fluxes about the previous time level. Here J denotes the Jacobian of the coordinate transformation. Dissipation operators, D_e and D_i are used in the central space differencing directions.

The smoothing terms used in the present study are of the form:

$$D_e|_\eta = (\Delta t)J^{-1}[\epsilon_2\bar{\delta}\rho(B)\beta\bar{\delta} + \epsilon_4\bar{\delta}\frac{\rho(B)}{1+\beta}\bar{\delta}^3]|_\eta J$$

$$D_i|_\eta = (\Delta t)J^{-1}[\epsilon_2\bar{\delta}\rho(B)\beta\bar{\delta} + 2.5\epsilon_4\bar{\delta}\rho(B)\bar{\delta}]|_\eta J$$

where $\beta = \frac{|\bar{\delta}^2 p|}{|(1+\bar{\delta}^2)p|}$ and where $\rho(B)$ is the true spectral radius of B . The idea here is that the fourth difference will be tuned down near shocks, that is, as β gets large the weight on the fourth difference drops down while the second difference tunes up.

For simplicity, all the boundary conditions have been imposed explicitly. On the body surface, the no-slip boundary condition is used and the wall temperature is specified. Free stream boundary conditions are used at the computational outer boundary. A symmetry boundary condition is imposed at the circumferential edges of the grid while a simple extrapolation is used at the downstream boundary. The flowfield is initially set to free stream conditions everywhere and then advanced in time until a steady state solution is obtained. Atmospheric flight conditions were used.

3. COMPOSITE GRID SCHEME

In the present work, a simple composite grid scheme⁸ has been used where a large single grid is split into a number of smaller grids so that computations can be performed on each of these grids separately. These grids use the available core memory one grid at a time, while the remaining grids are stored on an external disk storage device such as the solid state disk device (SSD) of the Cray X-MP/48 computer. The Cray-2 has a large incore memory to fit the large single grid. However, for accurate geometric modeling of complex projectile configurations which include blunt noses, sharp corners and base cavities, it is also desirable to split the large data base into a few smaller zones on Cray-2 as well.

The use of a composite grid scheme requires special care in storing and fetching the interface boundary data, i.e., the communication between the various zones. In the present scheme, there is a one to one mapping of the grid points at the interface boundaries and thus, no interpolations are required. Details of the data storage, data transfer and other pertinent information such as metric and differencing accuracy at the interfaces can be found in Reference 8 and 9. This scheme has been successfully used by Sahu⁹ to compute three dimensional transonic flow over two projectiles. The computed results clearly showed the transonic critical aerodynamic behavior in pitching moment coefficient observed in free flights. The present work is a further application of this technique to a more complicated projectile with base cavities.

III. MODEL GEOMETRY AND COMPUTATIONAL GRID

The external configuration of the M825, excluding the base, is similar to the M483A1 shown in Figure 1. The features of this projectile which have not been modeled exactly are the meplat on the fuze and the rotating band near the base. The rotating band was eliminated for simplicity and the meplat was modeled as a hemisphere cap. The computational model is shown in Figure 2 and consists of

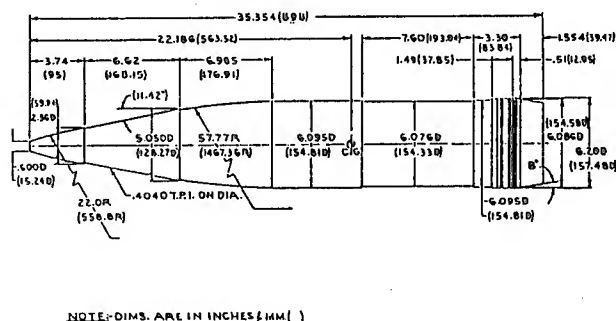


Figure 1. M825 Projectile

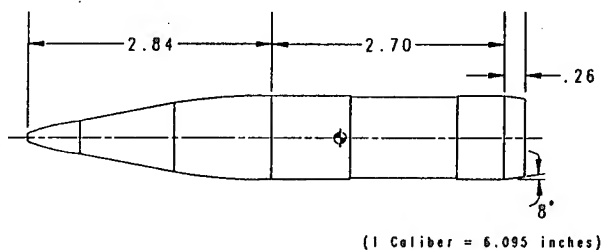


Figure 2. Computational Model

a 2.84 caliber nose, a 2.7 caliber cylindrical section, and a 0.26 caliber 8° boattail. The ogive contour as well as the undercut on the cylindrical section were matched.

The current problem of interest is the effect of the different base geometries on the overall projectile aerodynamics. Figure 3 shows the standard and dome base configurations. The standard base is a combination of aluminum and steel and contains a base cavity which is characterized as a flat surface. The PIP configuration is an all steel base and is characterized as a dome surface. The cavity volume is also significantly larger for the dome configuration.

The solution technique requires the discretization of the entire flow region of interest into a suitable computational grid. The grid outer boundary has been placed at 2.5 body lengths upstream and surrounding the projectile. The downstream boundary was placed at 2 body lengths. Since the calculations are in the subsonic/transonic regime the computational boundaries must extend out beyond the influence of the body. This ensures that the boundary conditions specified in the flow code are satisfied.

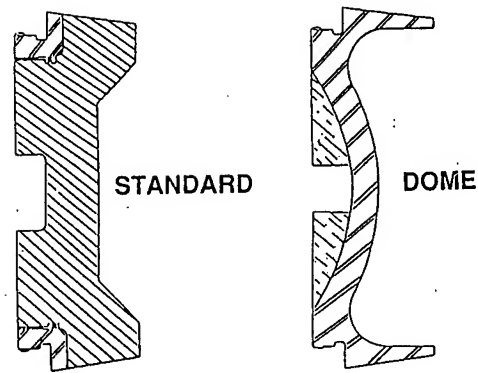


Figure 3. Base cavity configurations

Figure 4a and Figure 4b show the grids generated for the standard base and dome base configurations, respectively. Each of these grids consists of 225 points in the streamwise direction and 50 points in the normal direction. This is broken down into two sections: a body region and a base region. The surface points for each region are selected using an interactive design program. Each grid section is then computed separately using a hyperbolic grid generation program¹⁰. Longitudinally, there are 106 points along the projectile surface and 60 points in the base region downstream of the base corner. The normal distribution of points in base region consists of 50

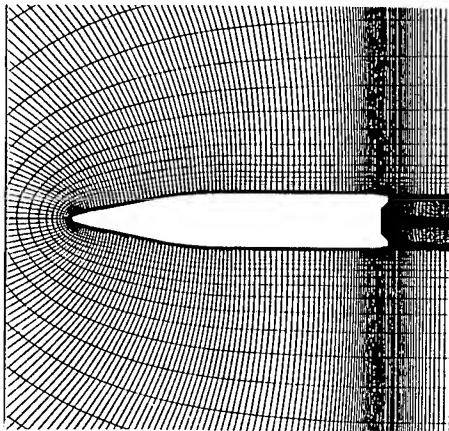


Figure 4a. Computational Grid, Standard Base

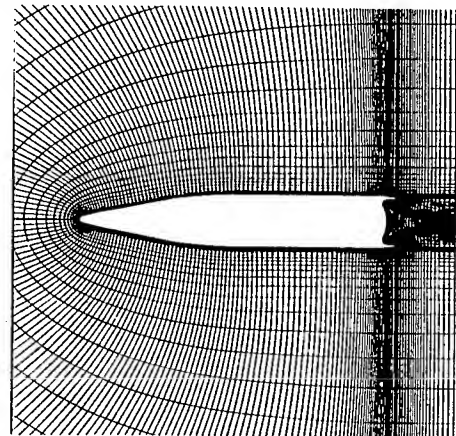


Figure 4b. Computational Grid, Dome Base

points along the base cavity. An expanded three dimensional view of the base grid is shown in Figure 5a. This grid has 33 points in the circumferential direction. The generally flat sections on the standard base enabled a grid to be routinely generated. However, due to the extreme concavity of the dome configuration, this grid (Figure 5b) required an increase in the smoothing values used by the hyperbolic grid generator, as well as the addition of a grid averaging technique.

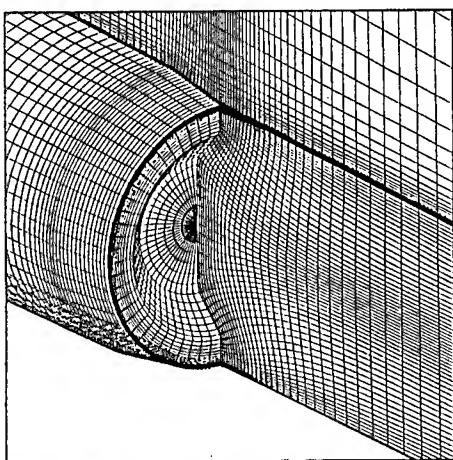


Figure 5a. Base Region Grid, Standard Base

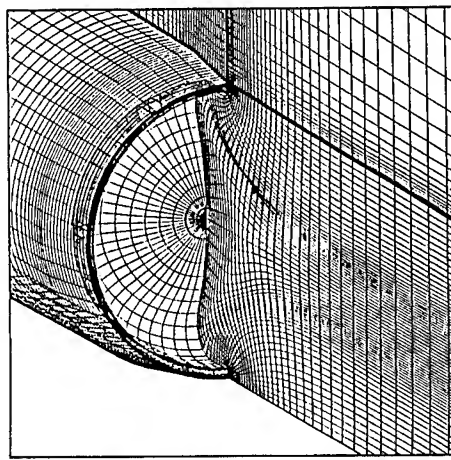


Figure 5b. Base Region Grid, Dome Base

IV. RESULTS

Numerical computations have been made for both the standard and the dome base configurations for a range of Mach numbers from $M = 0.80$ to 1.5 and at 4 degrees angle of attack. Computed results obtained at zero degree angle of attack are also included for comparison purposes.

A few qualitative results are presented next. Figure 6a and Figure 6b show the velocity vectors in the base region for both base configurations at $M_\infty = 0.98$ and $\alpha = 0.0^\circ$. The recirculatory flow in the base region is evident and as expected, is symmetric.

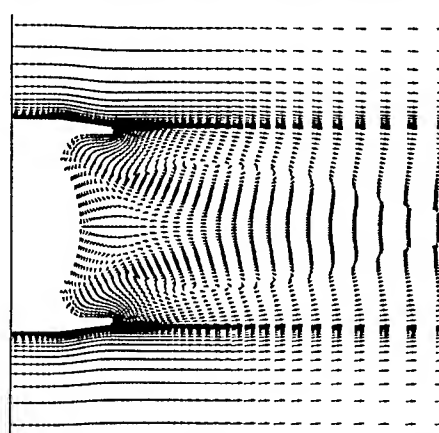
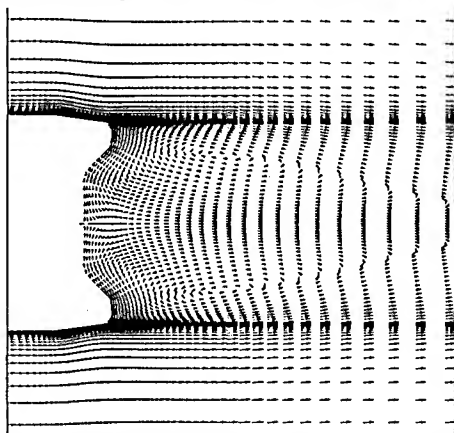


Fig. 6a Velocity vectors, $M = .98, \alpha = 0.0^\circ$ (Std) Fig. 6b Velocity vectors, $M = .98, \alpha = 0.0^\circ$ (Dome)

As shown in Figure 6a, the recirculation region for the standard base extends to about

one and a half caliber downstream of the base corner. The back flow, upon reaching the cavity follows the contour of the cavity and leaves the cavity pushing the flow upwards. The shear layer leaving the base corner is displaced upwards weakening the expansion at the base. Figure 6b for the dome configuration shows a weak secondary bubble inside the cavity in addition to the primary bubble. The flow again follows the contour of the cavity and, upon leaving the dome cavity, is almost parallel to the streamwise direction. This flow, thus, has less effect on the free shear layer and doesn't weaken the expansion at the base corner as much compared to the standard base. The net effect is that the size of the primary bubble for the dome base is slightly smaller than that for the standard base. The reattachment point is therefore closer to the base and results in lower base pressure or higher base drag at this Mach number⁵. Figure 7a and Figure 7b show the velocity vectors in the base region for the base configurations

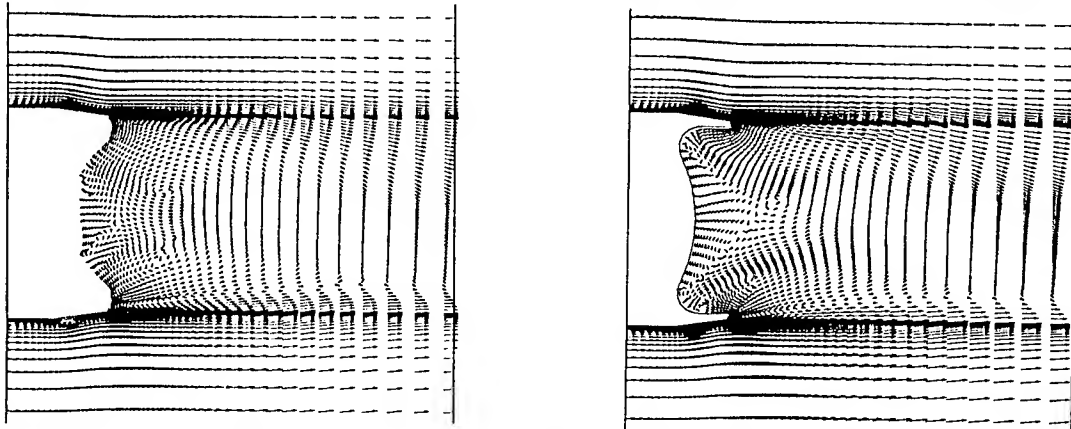


Fig. 7a Velocity vectors, $M_\infty = .98, \alpha = 4.0^\circ$ (Std) Fig. 7b Velocity vectors, $M_\infty = .98, \alpha = 4.0^\circ$ (Dome)

at $M_\infty = 0.98$ and $\alpha = 4.0^\circ$ for both windside (bottom half) and leeward side (top half). Again the recirculatory flow in the base region is evident and as expected, the flow in the wake is asymmetric. As shown in these figures, the separation bubbles on windside and leeward side in the wake differ in size and shape (the one on windside being more thin and elongated). In addition, a number of secondary separation bubbles can be seen to form inside the cavity for both base configurations. These changes in the flow structure contribute to small changes in the base pressure and thus, to the aerodynamic forces and moments.

Figure 8a and Figure 8b show the Mach number contours in the base region for both base configurations at $M_\infty = 0.98$ and $\alpha = 0.0^\circ$. These figures show the flow expansion at the ogive corner, boattail corner and the base corner. One can also see a shock wave on the cylinder portion of the projectile as well as a recompression shock system which exists downstream of the base corner. The flow field is symmetric for this condition. As angle of attack is increased to 4 degrees, the flow field becomes asymmetric (see Figure 9a and Figure 9b). A small asymmetry can be observed in the

location of the shock wave on the cylinder. The windside shock is further aft compared to the corresponding one on the leeside. The asymmetry can be clearly seen in the wake flow and its associated shock system. As can be seen in these figures, the wake flow field changes for the different base configurations.

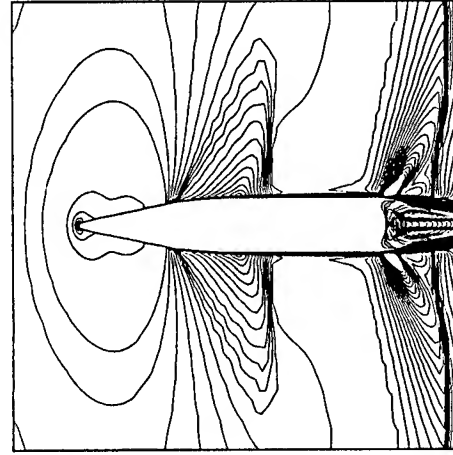
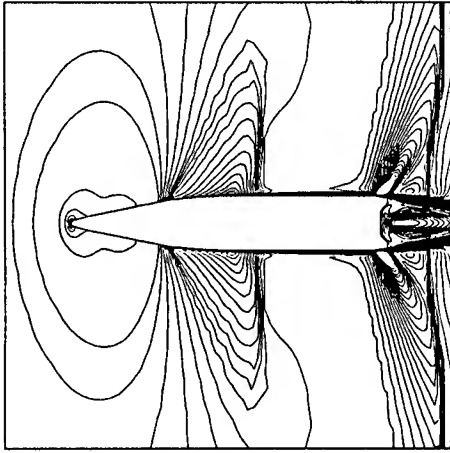


Fig. 8a Mach contours, $M_\infty = .98, \alpha = 0.0^\circ$ (Std) Fig. 8b Mach contours, $M_\infty = .98, \alpha = 0.0^\circ$ (Dome)

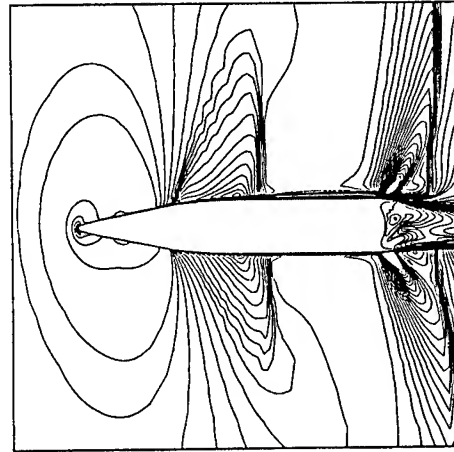
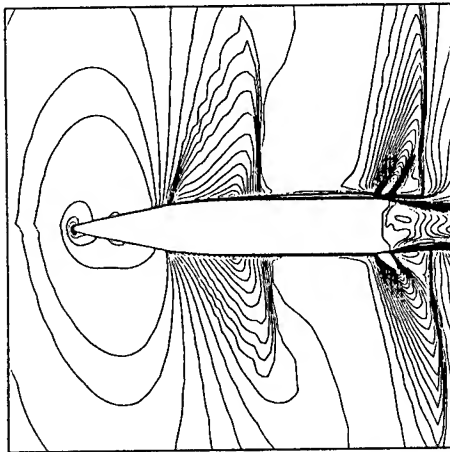


Fig. 9a Mach contours, $M = .98, \alpha = 4.0^\circ$ (Std) Fig. 9b Mach contours, $M = .98, \alpha = 4.0^\circ$ (Dome)

The entire flowfield over the projectile including the base region is computed. Therefore, the computed results include any upstream influence the base region flow may have on the boattail flowfield. Surface pressures including the base pressure and the viscous stresses are known from the computed flow field and can be integrated to give the aerodynamic forces and moments. Comparison of the zero yaw drag coefficient is shown in Figure 10. As seen in this figure, the difference in the computed drag is

small near $M_\infty = 0.97$. The computed results show higher drag for the dome base at $1.1 < M < 1.5$ and lower drag for $M < .92$. Also shown on this plot is the available range data for both configurations. The overall comparison of the computed drag with the range data is fair. As seen in the range data, the dome base has a higher drag, especially at higher transonic Mach numbers and this trend is seen in the computed results as well. Figure 11 shows the computed base pressure distribution for the dome base configuration at $M_\infty=1.1$ and $\alpha = 4.0^\circ$ for windside and leeside. As seen in this figure, the pressure on the windside ($Z/D=-0.5$) is higher than the pressure on lee-side. Since this pressure acts normal to the inside surface of the cavity, it produces a downward force (see Figure 12). Figure 12 shows the normal force coefficient for the dome base as a function of Mach number. The dotted line represents the normal force coefficient, C_N for the dome base projectile where the base region is excluded in the force and moment calculations. The solid line is for the entire projectile including the contribution from the base region. As mentioned

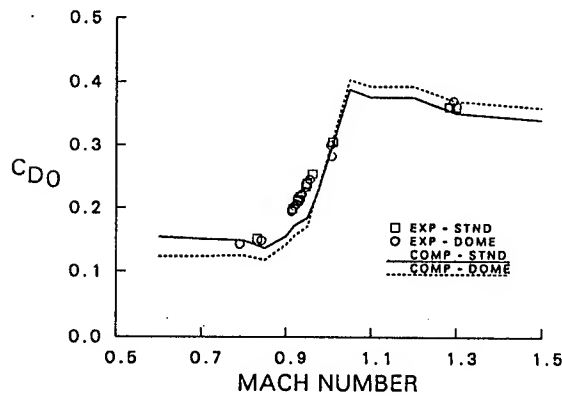


Figure 10 Total Drag vs Mach No

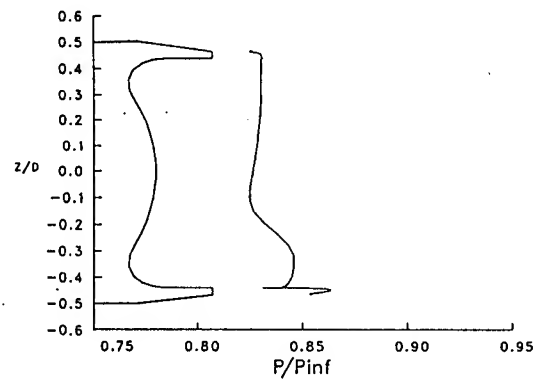
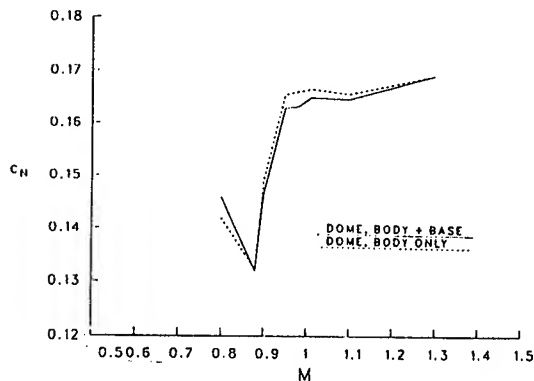
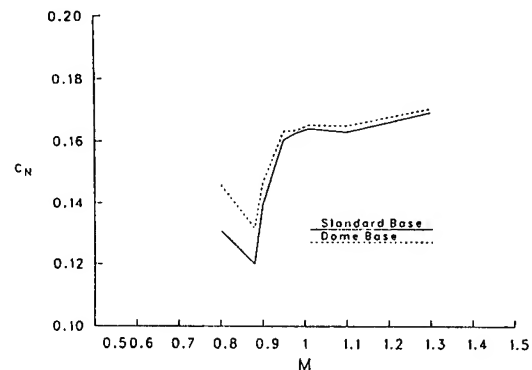
Figure 11 Base Pressure, $M_\infty = 1.1, \alpha = 4.0^\circ$ (Dome)

Figure 12 Normal Force vs Mach No

Figure 13 C_N vs Mach No (Std, Dome)

earlier, the base region produces a negative contribution and thus, the normal force is reduced somewhat for high transonic Mach numbers ($M > 0.90$). The reverse is true for low transonic speeds ($M < 0.85$). Figure 13 shows the normal force coefficient comparison for both base configurations as a function of Mach number. The dotted line represents the dome base result whereas the solid line shows the result for the standard base. As seen in this figure, the dome base has a higher normal force coefficient at transonic speeds compared to the standard base. The difference is small, of the order of a few percent, at higher transonic speeds ($M > 0.90$) and gets as large as 10-12 % at low transonic speeds ($M < 0.90$).

An aerodynamic coefficient which is of primary concern is the pitching moment coefficient, C_{m_α} . Figure 14 shows the C_{m_α} comparison for both base configurations. The computed C_{m_α} is also compared with the range data⁴ for both base configurations. Here C_{m_α} is referenced to the center of gravity of the projectile. The computed result clearly shows a sharp rise in C_{m_α} between $M = 0.80$ to 0.88 which is followed by a sharp drop as Mach number is increased to $M = 0.95$. As the Mach number is increased further C_{m_α} rises gradually again which is unlike the behavior of other projectiles such as the M549⁹. This critical behavior in C_{m_α} observed in the data is clearly predicted in the numerical computations. The overall comparison of the computed result with the range data is fair. As seen in the range data, the dome base configuration has lower C_{m_α} at transonic speeds compared to the standard base configuration and this trend is also seen in the computed results.

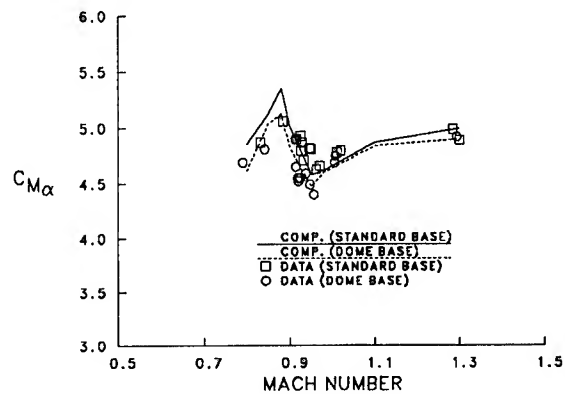


Figure 14 Pitching moment slope (Std, Dome)

V. CONCLUSIONS

Three dimensional numerical computations have been made for a projectile with two base cavity configurations at transonic speeds. Computed results show differences in the qualitative features of the base region flowfield between the two base cavities. Changes in the base cavity configuration have been found to affect the normal force and pitching moment coefficient. Differences in these coefficients of between 0 and 12% have been predicted and are compared with the range data. The dome base configuration produces a higher normal force and a lower pitching moment than the standard base at these transonic speeds and shows the same trend observed in the range data.

The use of CFD in this study has provided new insight to the fluid dynamic behavior responsible for small but important changes in the aerodynamics of the M825 projectile. This insight would have been impractical if not impossible to obtain through conventional testing means. The developed capability has proven to be very useful. Most importantly, it is available for design of future weapon systems and is currently being used to provide valuable design guidance in the development of an advanced multi-segment projectile system.

References

1. Sahu, J., Nietubicz, C.J. and Steger, J.L., "Navier-Stokes Computations of Projectile Base Flow with and without Base Injection," ARBRL-TR-02532, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, November 1983. (AD A125738) (also see AIAA Journal, Vol. 23, No. 9, September 1985, pp. 1348-1355.
2. Sahu, J., "Computations of Supersonic Flow over a Missile Afterbody Containing an Exhaust Jet," AIAA Journal of Spacecraft and Rockets, Vol. 24, No. 5, September- October 1987, pp. 403-410.
3. Sahu, J., "Three Dimensional Base Flow Calculation for a Projectile at Transonic Velocity," BRL-MR-3610, U.S. Army Ballistic research Laboratory, Aberdeen Proving Ground, Maryland, September 1987.
4. D'Amico, W.P., "Transonic Range Testing of the 155mm M825 Projectile with Standard and Dome Steel Bases," U.S. Army Ballistic research Laboratory, Aberdeen Proving Ground, Maryland, report to be published.
5. Sahu, J., Nietubicz, C.J. and Heavey, K.R., "Computational Study of the M825 Projectile with Standard and Dome Bases," U.S. Army Ballistic Research Laboratory, BRL-MR-3662, Aberdeen Proving Ground, Maryland, March 1988.
6. Pulliam, T.H. and Steger, J.L., "On Implicit Finite-Difference Simulations of Three-Dimensional Flow," AIAA Journal, Vol. 18, No. 2, February 1982, pp. 159-167.
7. Baldwin B.S. and Lomax, H., "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA Paper No. 78-257, January 1978.
8. Sahu, J. and Steger, J.L., "Numerical Simulation of Three Dimensional Transonic Flows," AIAA Paper No. 87-2293, Atmospheric Flight Mechanics Conference, Monterey, California, August 1987, (also see BRL-TR-2903, March 1988).
9. Sahu, J., "Numerical Computations of Transonic Critical Aerodynamic Behavior," Paper No. 88-4038-CP, AIAA/ASME/SIAM/APS 1st National Fluid Dynamics Congress, Cincinnati, Ohio, July 1988, (also see BRL-TR-2962, December 1988).
10. Nietubicz, C.J., Heavey, K.R. and Steger, J.L., "Grid Generation Technique for Projectile Configurations," ARO Report 82-3, Proceedings of the 1982 Army Numerical Analysis and Computers Conference, August 1982.

Finite Element Analysis of the Dynamic Thermal Buckling of a Thin Wall Cylinder

*Joseph M. Santiago, Dr., and Henry L. Wisniewski, Mr.
US Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

1 INTRODUCTION

The main objective of this investigation was to evaluate the accuracy of the finite element method, as embodied in the ADINA finite element program¹, in modeling the dynamic thermal buckling of a thin wall cylindrical shell. Accuracy was measured by comparison with data from an experiment in which one side of an aluminum cylinder was subjected to a 2 second duration thermal pulse simulating exposure to a nuclear explosion. A subsidiary objective was to maximize the computational efficiency of the calculations by judiciously adjusting the size of the time step to suit the different phases of the buckling process.

A major problem in calculating the dynamic buckling of a structure is the different time scales that characterize the buckling process. The initial loading phase is essentially quasi-static and of long duration. The unloading phase (often referred to as *dynamic snap-through*) that occurs immediately after reaching the critical buckling load is extremely dynamic, nonlinear, and of relatively short duration. The final phase, again of long duration, is characterized by gradually diminishing oscillations as the deformation changes from a moderately dynamic to a quasi-static state.

What is most interesting about the present analysis is the unique way this last problem is dealt with by using an implicit time integrating scheme. By varying the size of the time step in the integrating scheme the amount of inertia introduced into the equations of motion can be controlled. Thus, the relatively large time step used during the initial loading phase virtually eliminated inertial effects to produce a quasi-static solution. The very small time step used during the sudden unloading phase automatically restored the appropriate inertial terms to produce a correct dynamic solution. The incrementally increasing time step used during the final oscillatory phases gradually diminished inertial effects to permit a smooth transition from a dynamic to a quasi-static solution. In this way the entire calculation was performed without interruption using a single numerical integration algorithm in which the size of the time step was adjusted to optimally

model the different buckling phases. Thus, the total number of computational cycles was minimized without compromising accuracy, resulting in the total cpu time being reduced.

This paper continues a computational effort to reproduce the thermally induced buckling deformation measured in the second of a series of tests in which a thin wall cylinder was exposed to intense thermal radiation². The first computations evaluated the linearized buckling analysis method and the nonlinear quasi-static analysis method and found that both methods correctly predicted buckling near the fixed ends of the cylinder³. These were followed by a dynamic nonlinear analysis to model the post-buckling deformation, but without the use of an initial imperfection, relying on the spatially nonuniform thermal loading to induce buckling⁴. The results were found to agree strikingly well with test results, but were subsequently found to have an error in element connectivity that inadvertently acted as a buckling initiating imperfection. This paper reports on subsequent calculations to reproduce test results with this error corrected.

2 DESCRIPTION OF PROBLEM

The cylinder analyzed in this paper has the following dimensions:

$$\begin{aligned} L &= \text{length between clamped ends} = 0.8 \text{ m} \\ D &= \text{inside diameter} = 0.3048 \text{ m} \\ h &= \text{wall thickness} = 1.016 \text{ mm} \end{aligned}$$

The mass density was taken as the nominal value for aluminum at 27 °C, reference 5:

$$\rho = \text{mass density} = 2700 \text{ kg/m}^3$$

The temperature dependent properties for the 6061-T6 aluminum, linearized between the limits of 0°C to 250°C for use in the previous investigation⁴, are summarized in Table 1.

Table 1: Linear fit to the temperature dependent mechanical properties of 6061-T6 aluminum

T	temperature	0 °C	250 °C
E	Young's modulus	74.19 GPa	64.70 GPa
ν	Poisson's ratio	0.3285	0.3285
σ_y	yield stress	307.6 MPa	252.6 MPa
E_T	plastic tangent modulus	2368 MPa	872.2 MPa
α	thermal expansion coefficient	$22.2 \times 10^{-6}/^\circ\text{C}$	$27.2 \times 10^{-6}/^\circ\text{C}$

The cylinder was instrumented with thermal couples at a number of locations to record temperatures. Spatially, the temperature was found not to vary substantially

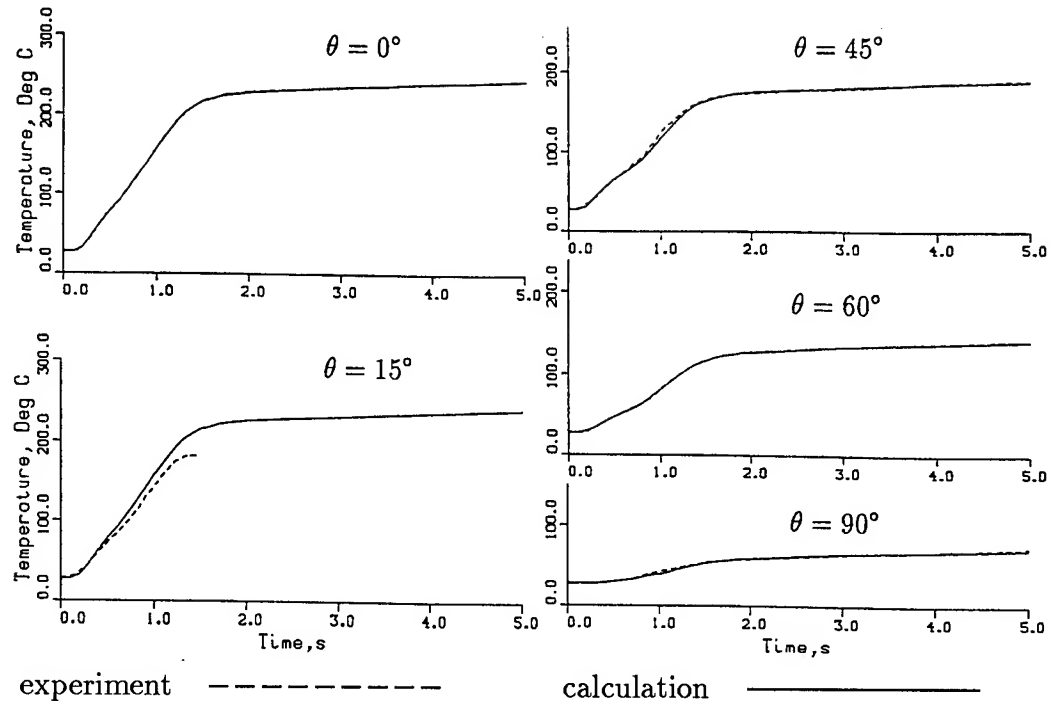


Figure 1: Comparison of experimental temperature histories recorded at 0°, 15°, 45°, 60°, and 90°, with loading functions generated for the ADINA analysis.

in the axial direction, so that the temperature histories recorded on the circumference midway between the cylinder ends, shown as dash lines in Figure 1 with θ measured circumferentially from the location closest to the source, were sufficient to characterize the temperature loading. These records were extended to provide temperature histories at all nodes of the finite element model using the relation

$$T = \frac{T_0 - T_{180}}{1 + \alpha^2 \tan^{2n} \frac{1}{2}\theta} + T_{180} \quad (1)$$

where T_0 is the temperature recorded at 0°, as shown in Figure 1, and T_{180} is the temperature at 180°, set equal to the recorded ambient temperature of 27.78 °C. α and n were adjusted to make the temperature agree with the temperature recorded at 60° and to approximate as closely as possible the temperatures recorded at 45° and at 90°. Spatially this equation gave a temperature distribution that decreased monotonically from a maximum at 0° to a minimum at 180°.

Although in tests involving substantial temperature changes, strain gages corrected to match the thermal expansion of the specimen are usually employed, so that they record strains due only to stresses and not those due to thermal expansion, these were

not available for the present test and consequently the strain records had to be corrected after the test using data from a subsequent free thermal expansion furnace test². Since the ADINA program computes the sum of the stress induced strains and the thermal expansion strains, the computed strains had to be similarly corrected before they could be compared with the experimental strains.

3 FINITE ELEMENT MODEL

Advantage was taken of the spatial symmetry of the geometry and the temperature distribution to model the problem using one quarter of the cylinder, as shown in Figure 2. A model consisting of 703 nodes and 72 elements was generated using 16 node shell element resident in ADINA, reference 1, Section XVII. The fixed boundary conditions on the cylinder end and symmetry conditions on the symmetry planes resulted in a 3184 degrees-of-freedom model. The calculations used a $3 \times 3 \times 2$ Gauss point order of integration, found from previous studies to be most accurate⁶. The temperature dependent material properties tabulated in Table 1 were input into the ADINA linear kinematic hardening material model, reference 1, pages XVII.47-52.

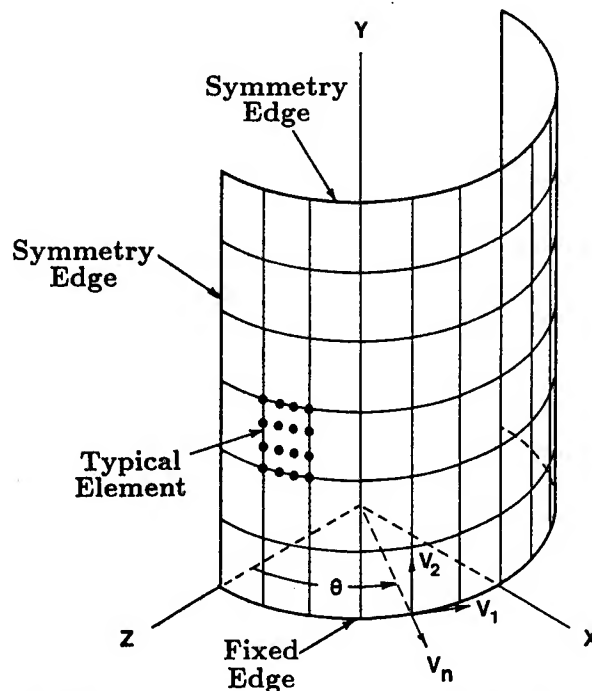


Figure 2: Finite element model of the quarter cylinder showing the element mesh.

4 LINEARIZED BUCKLING ANALYSIS

Briefly, the linearized buckling analysis method⁷ is based on linearly extrapolating the stiffness matrix and the associated load intensity from two distinct load steps to obtain an estimate on the critical load. Buckling is predicted when the extrapolated matrix becomes singular, resulting in an eigenvalue problem on the extrapolated value of the load intensity. The linearized buckling algorithm solves the eigenvalue problem and then uses the resulting singular stiffness matrix to determine the corresponding eigenvector, which represents the buckled displacement field or buckling mode. The critical load predicted by this method will typically exceed the true buckling load, but this is inconsequential since the primary purpose of the method is to generate buckling modes for subsequent use in the nonlinear analysis as initial geometric imperfections to induce buckling.

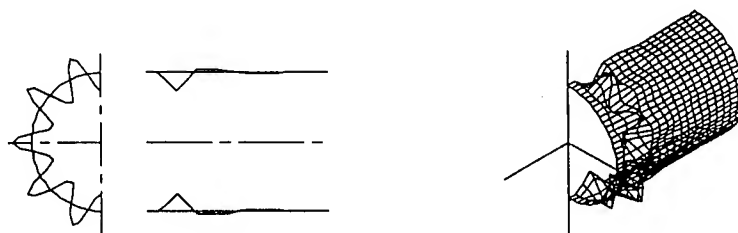


Figure 3: First thermal buckling mode of the cylinder as calculated by the ADINA linearized buckling analysis. Deflections magnified 300 times.

For the problem treated here, rather than using the experimental temperatures, the linearized analysis employed for simplicity a spatially uniform temperature distribution increasing in one step from 27.78 °C to 28.78 °C. Solutions were obtained for the first 10 buckling modes, of which the first, shown in Figure 3, is typical. The isometric view in the figure has the coordinate axes located at the fixed end, the first cross-section is perpendicular to the cylinder axis at a location 44.45 mm from the fixed end, and the second cross-section is parallel to the axis through the crown line. Buckling for the lower modes was primarily confined to the regions near the fixed ends with well defined circumferential wave patterns, in agreement with test measurements.

5 NONLINEAR BUCKLING ANALYSES

Three different nonlinear calculations were performed. The first assumed that the cylinder was geometrically perfect and depended on the spatial nonuniformity of the temperature loading to induce buckling. The second and third calculations assumed

the cylinder to have initial geometric imperfections proportional to the first linearized buckling mode, Figure 3. The second calculation represented a moderate imperfection corresponding to a maximum radial deviation equal to 10% of the wall thickness. The third calculation assumed a maximum radial deviation of 50% of the wall thickness, more in line with the maximum geometric irregularities measured on the specimen, reference 2, Figures 26 and 27. All three calculations used the temperature loading functions generated by equation (1).

Table 2 shows the time step schedule developed after some preliminary trials for the first (perfect geometry) calculation involving a total of 2280 time steps for 5 seconds of problem time. We observe that the moderately large steps used during the initial loading phase are gradually transformed to a minimum size shortly before buckling and after buckling gradually return to their former large size. Note that it was again necessary to diminish the time step in the interval from 1126.4 to 1200 ms because of numerical difficulties encountered there in converging. Similar schedules to synchronize step size with buckling phase were developed for the second and third calculations, which involved totals of 1606 and 1666 steps, respectively, but these are not shown due to space limitations.

Table 2: Intervals of time steps used in the ADINA dynamic buckling calculation.

Time blocks	Step size (ms)	Number of steps	Start time (ms)	End time (ms)
1	100.00	1	0	100
2	10.00	21	100	310
3	9.00	7	310	373
4	8.00	28	373	597
5	4.00	26	597	701
6	3.00	31	701	794
7	2.00	53	794	900
8	0.80	50	900	940
9	0.40	30	940	952
10	0.05	60	952	955
11	0.10	22	955	957.2
12	0.40	423	957.2	1126.4
13	0.20	368	1126.4	1200
14	0.50	600	1200	1500
15	1.00	500	1500	2000
16	50.00	60	2000	5000

The Newmark time marching scheme was used to integrate the equations of motion, reference 1, page II.53. In order to provide numerical dissipation values of

$$\delta = 0.7 \quad \text{and} \quad \alpha = 0.4$$

were used for the Newmark parameters, rather than the default values. Implemented in this way, dissipation is made directly proportional to step size⁸. Consequently, not only were transitions between step size smoothed, but also little dissipation was introduced into the dynamic snap-through phase to maintain accuracy, while dissipation was gradually increased during the final post-buckling phase to reduce oscillations. In agreement with the use of an implicit analysis method, a consistent mass formulation was used, reference 1, page II.38. The total Lagrangian method of nonlinear analysis was employed, reference 1, page XVII.14, and the solution was performed using the full Newton method, with equilibrium iterations at each time step, reference 1, page II.53.

6 RESULTS OF NONLINEAR BUCKLING ANALYSES

The three problems were run on the Cray X-MP/48 computer at the Ballistic Research Laboratory. The cpu time required for each problem ranged from a minimum of slightly less than 13 hours to a maximum of 18.5 hours, the difference primarily reflecting the total number of time steps required for each calculation.

1. No Initial Imperfection

Figures 4-6 summarize the results of the nonlinear analysis assuming an initially perfect cylinder. Figure 4 shows the deforming cylinder just before buckling, just after buckling, and at the end of the calculation (5 seconds), the orientation and cross-sectional views being the same as in Figure 3. Before buckling the surface of the cylinder facing the heat source undergoes a small uniform radial expansion due to heating, except for a small kink beginning to form near the fixed end. Immediately after buckling the displacements have dramatically increased. Longitudinally, there is an outwardly bulged region next to the fixed end, followed by a inwardly kinked region, which merges into a second bulged region at the center. Circumferentially, there are three well defined lobes. At the end of the calculation the continued heating has resulted in an intensification of the deformation immediately after buckling, but with no essential change in the overall pattern. This deformation, however, does not agree with post-test measurements, which show *four* rather than *three* circumferential lobes and the region next to the fixed end inwardly kinked rather than outwardly bulged, reference 2, Figures 25 and 27.

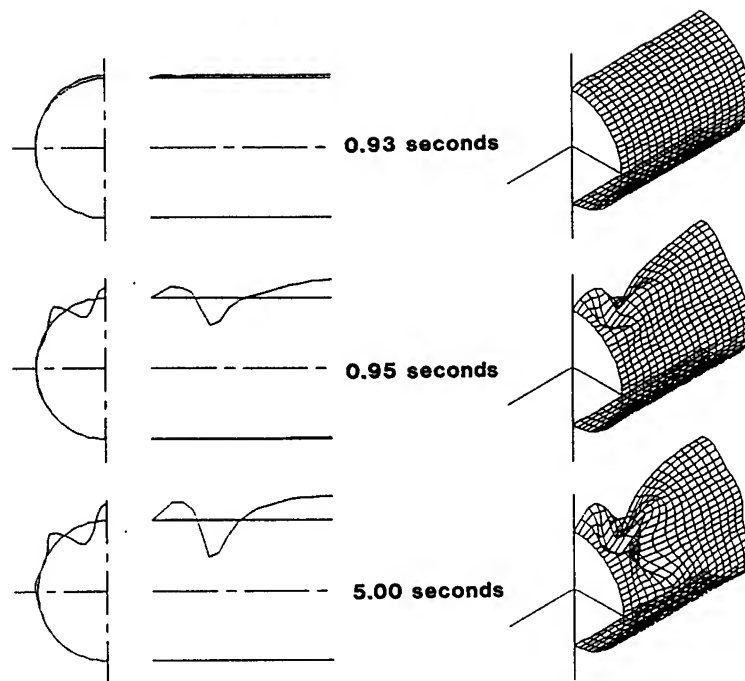


Figure 4: Calculated deformations with no initial imperfection. Deflections magnified 10 times.

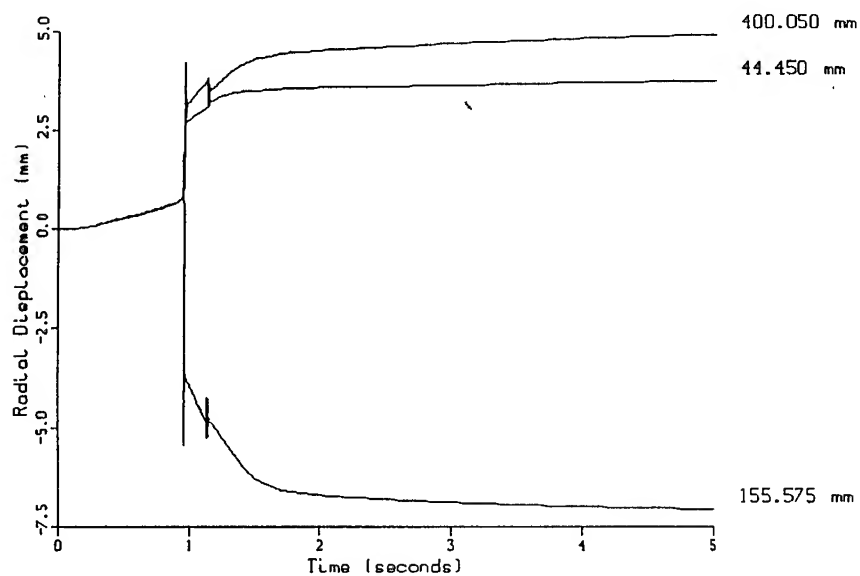


Figure 5: Calculated histories of the radial deflection at three points on the crown line assuming no initial imperfection.

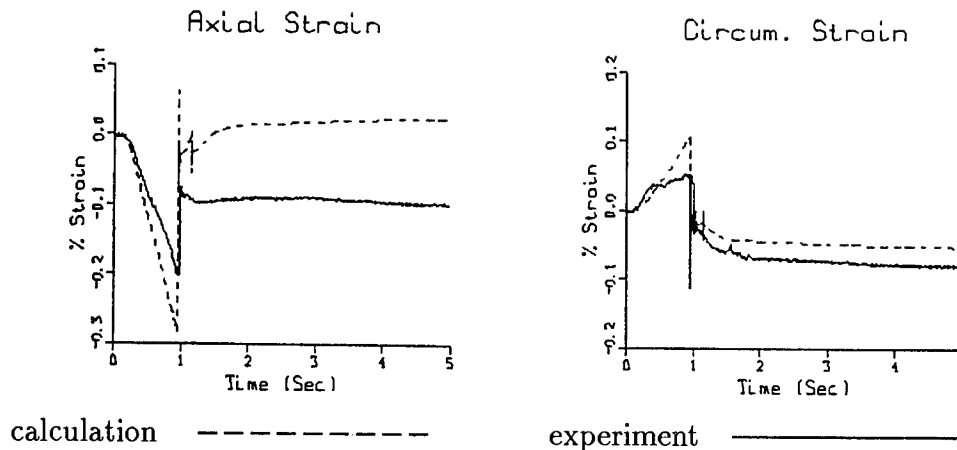


Figure 6: Comparison of the thermally corrected axial and circumferential strain histories at the center of the crown line calculated assuming no initial imperfection.

Figure 5 shows the histories of the radial displacement at three points along the crown at a distance of 44.45 mm, 155.575 mm, and 400.05 mm from the fixed end, the latter coinciding with the middle of the cylinder. It shows that buckling is first predicted to occur at 0.95 seconds with a smaller jump at 11.5 seconds, the first being remarkably close to the 1 second detected in the test records.

Figure 6 compares the calculated and the experimental thermally corrected axial and circumferential strain histories, their location being at the center of the crown line on the inner surface of the cylinder. The comparison shows that the calculated time of buckling corresponds closely with that recorded experimentally. Moreover, the calculated histories show the same characteristic as the experimental records with the circumferential strains being quantitatively closer.

2. 10% Initial Imperfection

Figures 7-9 summarize the results of the nonlinear analysis for the 10% initial imperfection. Figure 7 is similar to Figure 4, showing the deforming cylinder just before buckling, just after buckling, and at the end of the calculation. The deformations show the same general behavior as before, a small uniform radial expansion of the front surface with a small kink near the fixed end before buckling, a sharp nonuniform increase in the deflections immediately after buckling, and an intensification of the buckled deformation at the end of the calculation. However, the pattern of the deformation immediately after buckling is noticeably different from that calculated before. Longitudinally, the region adjacent to the fixed end is now inwardly kinked and there is a gradual merger to a

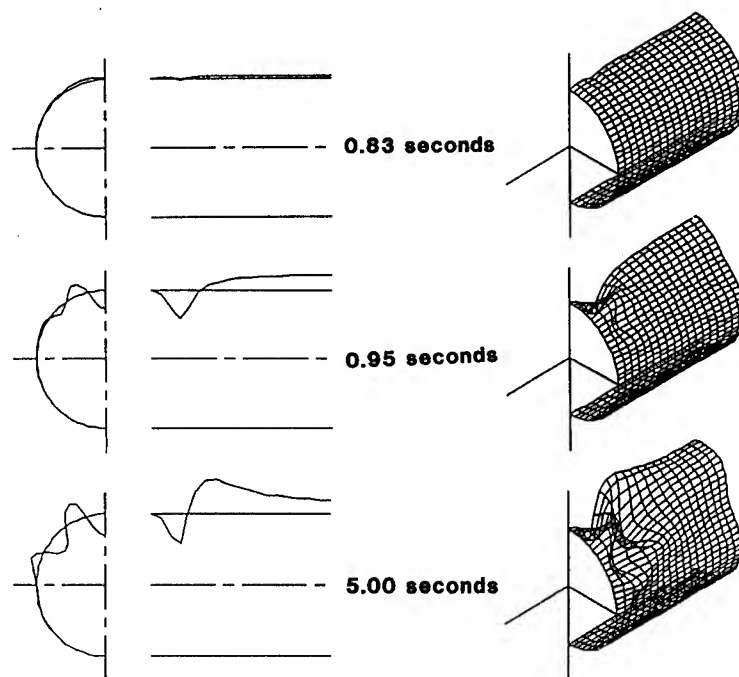


Figure 7: Calculated deformations with a 10% initial imperfection. Deflections magnified 10 times.

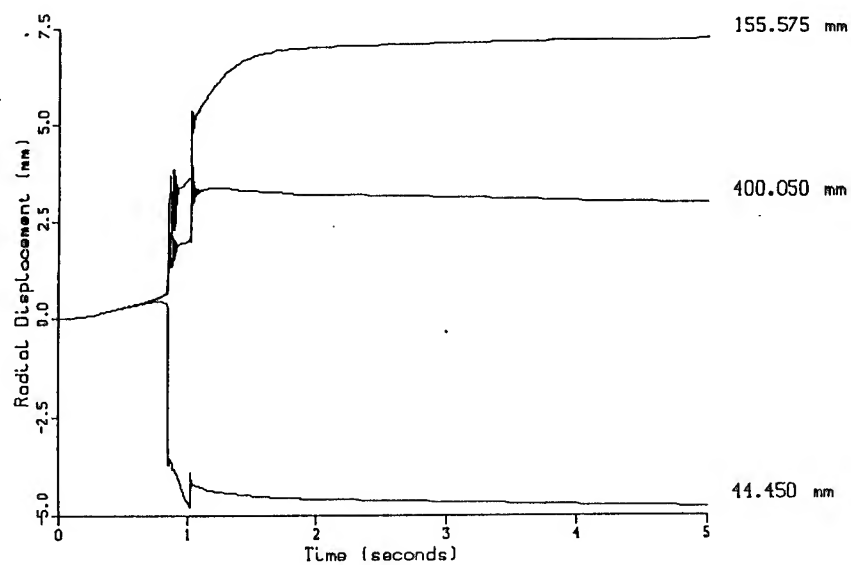


Figure 8: Calculated histories of the radial deflection at three points on the crown line assuming a 10% initial imperfection.

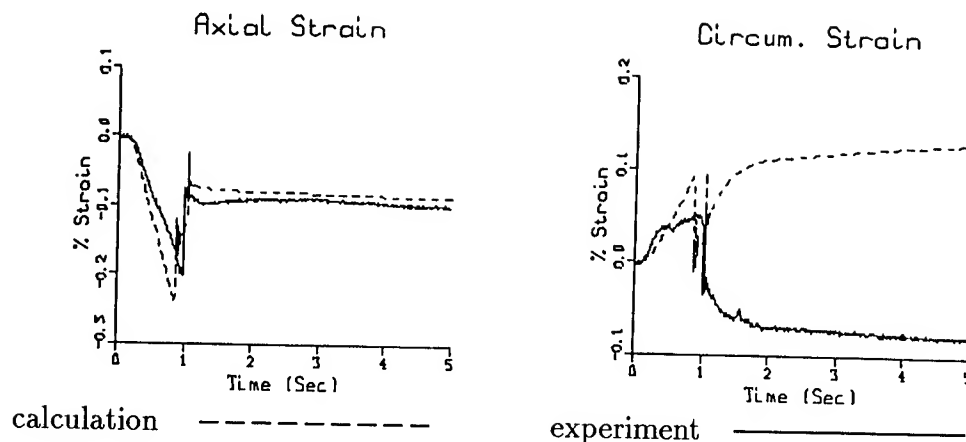


Figure 9: Comparison of the thermally corrected axial and circumferential strain histories at the center of the crown line calculated assuming a 10% initial imperfection.

single outwardly bulged region at the center. Circumferentially, there is a well defined *four* lobed pattern, more closely resembling the experimental deformation. The last view shows that the final deformation is not simply an intensification of the intermediate deformation, but that the region of maximum radial expansion has shifted from the center to the region next to the inward kink.

Figure 8 shows the displacement histories at the same locations as in Figure 5. As in the previous calculation, buckling again occurred in two distinct jumps, this time at 0.85 seconds and 10.5 seconds, with the second being closer to that detected on test records. The curves also explain the shift in the region of maximum expansion observed in in Figure 7, showing it is caused by the deflection at 155.575 mm suddenly increasing at the second jump from a value less than that at 400.050 mm to a value greater. Figure 9 is similar to Figure 6, comparing the calculated and the experimental strain histories. It shows that there is now a remarkable coincidence in the axial strain histories, while the discrepancy in the late time histories of the circumferential strains has considerably worsened.

3. 50% Initial Imperfection

The results of the ADINA calculation for a 50% initial imperfection are summarized in Figures 10–12. Like Figures 4 and 7, Figure 10 shows the deformation before buckling, after buckling, and at the end of the calculation. The deformations after buckling and at the end of the calculation resemble very much the final deformation of the 10% imperfection calculation. Four circumferential lobes are again apparent, but not as distinctly

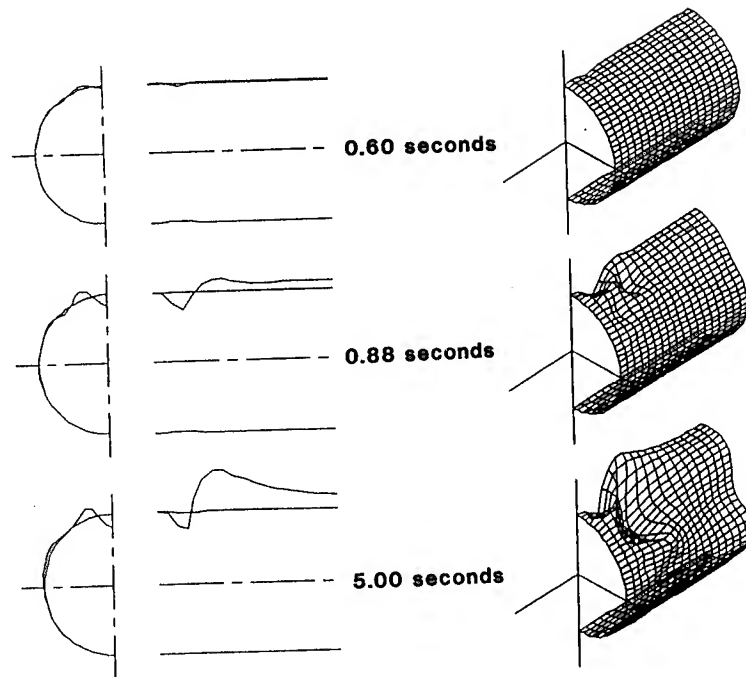


Figure 10: Calculated deformations with a 50% initial imperfection. Deflections magnified 10 times.

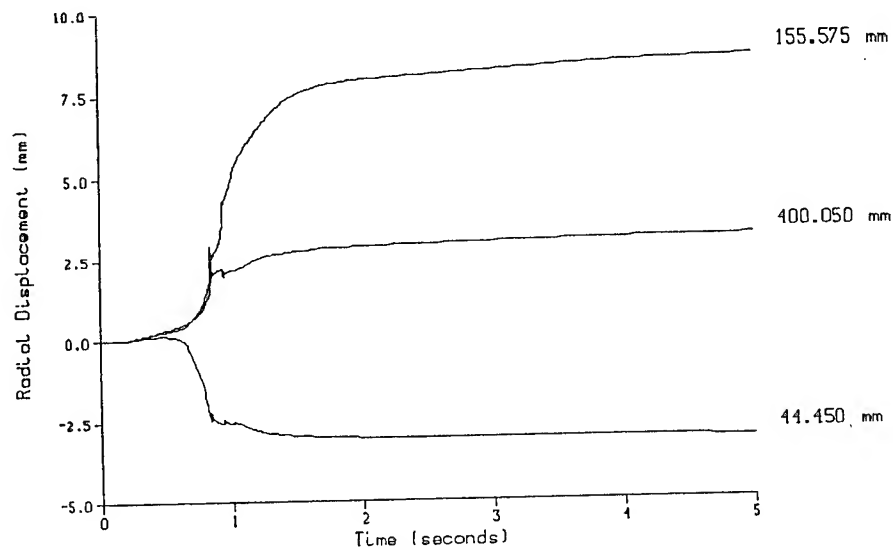


Figure 11: Calculated histories of the radial deflection at three points on the crown line assuming a 50% initial imperfection.

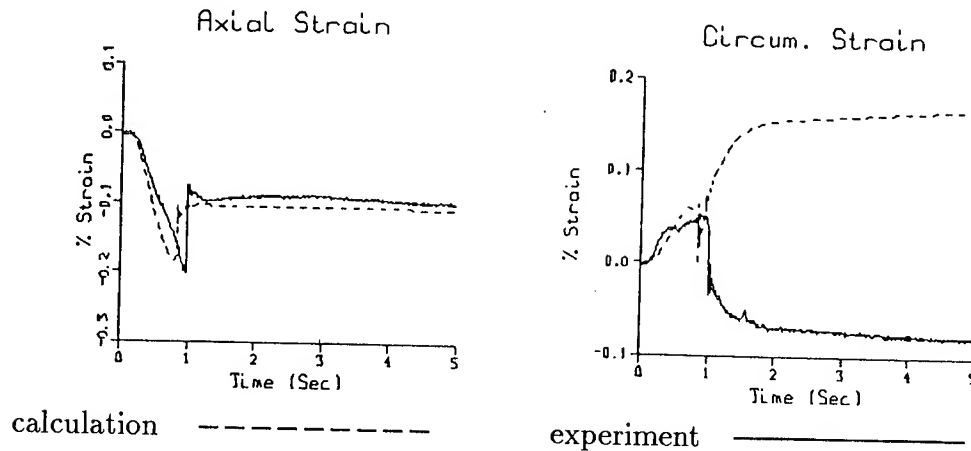


Figure 12: Comparison of the thermally corrected axial and circumferential strain histories at the center of the crown line calculated assuming a 50% initial imperfection.

as in the previous calculation. The last view shows that, as before, the final deformation resembles the intermediate deformation with the expected intensification.

Figure 11 is like Figures 5 and 8, showing the displacement histories at the same locations. Although the time of buckling can be localized to 0.85 seconds, the demarcation between no buckling and buckling has been considerably blurred by the severity of the initial imperfection. Moreover, the divergence of the deflections from a uniform radial expansion is noticeable much earlier and it is difficult to detect a secondary buckling. Figure 12 shows that the close correlation between the computed and experimental axial strain histories is as good as that computed with a 10% imperfection, Figure 9. On the other hand, the discrepancy in the late time histories of the circumferential strain has become even worse.

7 SUMMARY OF RESULTS

All three calculations captured the salient features of the thermal buckling of the cylinder between fixed ends. They all predicted buckling of the front face of the cylinder near the fixed ends and showed little deformation of the rear face. Initially the calculations predict a nearly uniform radial deflection of the front face resulting from circumferential expansion due to heating and the Poisson's effect of the cylinder being constrained from expanding in the axial direction. The initial build-up of the compressive strains in the axial direction confirms the constraining effects of the fixed ends.

The deformation predicted with no initial imperfection did not match the final deformation measured in the thermal radiation test. With an imperfection of either 10% or 50% of the wall thickness, the same overall deformation as measured is predicted, confirming the sensitivity of the final configuration to imperfections. With no imperfection, buckling was calculated to occur first at 0.95 seconds followed by secondary buckling at 1.15 seconds. With a 10% imperfection, primary buckling occurred at 0.85 seconds and secondary buckling at 1.05 seconds. With a 50% imperfection, only primary buckling was predicted at 0.85 seconds. They all closely straddle the experimentally recorded buckling time of 1 second and confirm the accuracy of the calculations, implying that initiation of buckling is controlled primarily by the amount of heating, *i.e.*, by the load intensity, rather than the degree of imperfection. The comparison of the strain histories show that all three calculations are relatively accurate in predicting strains before buckling. After buckling the no imperfection calculation predicts the circumferential strain best, while the two imperfection calculations are superior in predicting the axial strain. Although this is puzzling, it is still remarkable how well the computed strains reproduced the salient features of the recorded strains, given the well known difficulties in numerically calculating strains accurately. Lastly, both the deflection and strain history plots verify the effectiveness of the dissipative time integration scheme in reducing oscillations after buckling and smoothing transitions between changes in time step sizes.

8 CONCLUSIONS

The ADINA nonlinear finite element analysis method using the shell element model with thermal stress capability can predict thermal buckling accurately. To reproduce the thermal buckling measured in the thermal radiation test reasonably well it was necessary to use an initial geometric imperfection. The first buckling mode generated by the linearized buckling analysis method was found to work well as an initial imperfection. The calculations predicted the instant of buckling accurately, even without an initial imperfection. The calculated strain histories reproduce the salient features of the test records sufficiently well and the correlation can be considered successful, given that strains are typically more difficult to calculate accurately than are displacements. Although each nonlinear analysis was computer intensive, the use of a variable time step size contributed significantly to the efficiency of the calculation, solving 5 seconds of a highly nonlinear dynamic buckling problem in a total of 1606 to 2280 time steps. Also, the numerical dissipation introduced through the time integrating scheme proved very effective in reducing oscillations and smoothing transitions between time steps.

9 ACKNOWLEDGMENTS

We gratefully acknowledge the Nuclear Weapon Effects Project Office of the US Army Harry Diamond Laboratory for funding the investigation.

10 REFERENCES

1. ADINA, *A Finite Element Program for Automatic Dynamic Incremental Nonlinear Analysis, Users Manual*, Volumes 1 & 2, Report ARD 87-1, ADINA R & D, Watertown, Massachusetts, December 1987.
2. R. J. PEARSON, H. L. WISNIEWSKI, P. D. SZABADOS, and A. W. WILSON, *The Effects of Thermal/Blast Synergism on the Nuclear Vulnerability of a Generic Aircraft Structure*, Technical Report ARBRL-TR-02540, US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, January 1984.
3. C. R. MATISZIK, J. M. SANTIAGO, and H. L. WISNIEWSKI, "Thermal Buckling of a Thin Cylindrical Shell," *Proc. 6th Intrn. Conf. Numerical Methods in Thermal Problems*, Swansea, UK, July 1989.
4. J. M. SANTIAGO, C. R. MATISZIK, and H. L. WISNIEWSKI, "Dynamic Thermal Buckling of a Thin Cylindrical Shell," to be published in the *Proc. 11th Intrn. MABS Symposium*, Albuquerque, NM, September 1989.
5. *Engineering Data For Aluminum Structures, Aluminum Construction Manual*, The Aluminum Association, 750 Third Avenue, New York, NY 10017, August 1970.
6. J. M. SANTIAGO and H. L. WISNIEWSKI, "Convergence of finite element frequency predictions for a thin walled cylinder," *Computers & Structures*, **32**, 745-759 (1989).
7. K. J. BATHE and E. DVORKIN, "On the Automatic Solution of Nonlinear Finite Element Equations," *Computers & Structures*, **17**, 871-879 (1983).
8. F. G. LATURELLE, "Finite element analysis of wave propagation in an elastic half-space under step loading," *Computers & Structures*, **32**, 721-735 (1989).

Spectral Luminescence of Camouflage Fabrics and Paints

Melvin B. Satterwhite, Mr.
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, VA 22060-5546

The detection of military targets has been successful when using high spatial and spectral resolution remote sensing technologies because some present camouflage fabrics and paints are ineffective in emulating soil or vegetation. Other camouflages are excellent emulators of vegetation or soil backgrounds. Future improvements in camouflage materials should make them more effective, in which case new remote sensing technologies will be needed. Laboratory measurements of present camouflage materials, vegetation, and soil have addressed the hyperspectral and thermal characteristics of these materials,¹ but the spectral luminescence features of these materials have not been widely studied.²

Spectral luminescence is the process by which excitation photons impinging upon a material bring about the emission of photons at a longer wavelength. This process includes both fluorescence and phosphorescence which describe time dependent photon emissions.³ Some field studies have attempted to use luminescence remote sensing for mineral exploration, but with little success.⁴ This may have resulted from inadequate understanding of the luminescence characteristics of terrain features or from selecting inappropriate image wavebands for analysis. Even so, laboratory studies of mineral luminescence have shown the feasibility of using spectral luminescence for geological investigations,^{5,6} as well as for terrestrial vegetation studies.^{7,8}

Passive remote sensing luminescence in the terrestrial environment has been conducted using the Fraunhofer Line Discriminator (FLD).⁹ This approach is based on the Fraunhofer line depth principle, and uses solar-stimulated luminescence in selected Fraunhofer absorption lines. The airborne FLD measurements have validated the concept of acquiring remotely-sensed luminescence data and imagery, but successful image exploitation requires additional work.⁴

An active remote sensing approach could use laser-induced luminescence (LIF). This technology may have wider application because it could be used

at night and there would be greater flexibility in selecting excitation and emission wavebands. Specific excitation-emission waveband combinations could be selected for optimal materiel discrimination rather than being confined to the Fraunhofer or atmospheric absorption lines. The system would not be hindered by reflected sunlight nor the solar-stimulated background luminescence.

Although luminescence characteristics of some minerals⁶ and vegetation^{7,8} have been described, little has been done to describe the luminescence features of military camouflage fabrics or paints.² The paucity of luminescence data for man-made materials limits the application and exploitation of luminescence remote sensing technology and creates uncertainties with respect to the selection of optimal excitation-emission wavebands for target detection.

OBJECTIVE

The objective of this laboratory study is to determine the effectiveness of camouflage fabrics and paints to emulate the luminescence of vegetation and soil backgrounds.

PROCEDURES AND MATERIALS

Samples of garnish fabrics used in making camouflage screens, plastics with or without an embedded fiber netting, and fabrics made from woven or felted natural plant fibers were obtained. The garnish color was determined by visual analysis. Samples of camouflage paint were measured. Each paint sample had been sprayed on a 2 x 2 cm aluminum plate that had been previously sprayed with a reddish-brown-colored primer paint.

Soil samples were obtained and a subsample of the air-dried, <250 μ m sieved separate was measured. The vegetation samples were leaves or leaflets selected for differences in leaf type, leaf moisture content, and leaf senescence stage. The leaf types were broadleaf, needle leaf, and graminoid. Leaf moisture conditions were (1) the freshly picked turgid green leaf, and (2) the air-dried green leaf. Leaf senescence was determined by visual assessment of the leaf color; green, yellow-green, green-yellow, yellow, or brown. All granular samples were placed directly into powder sample holders fitted with a fused silica glass window. All samples were measured in the front surface mode at room temperature.

Spectral luminescence measurements were made using a SLM Instruments spectrofluorometer, model 8000C. In the normal operation mode, the excitation monochromator was held constant and a narrow, 4 nm, waveband of light irradiated a 0.3 x 1 cm spot on the sample. The excitation monochromator was stepped at 10 nm intervals over the 250-800 nm spectrum, and the emission monochromator was stepped at 5 nm intervals over the 300-850 nm

spectrum. The starting wavelength of an emission spectrum was 20 nm greater than the associated excitation wavelength. Emissions from the sample surface were directed in a narrow solid angle to the emission monochromator and onto the emission photomultiplier tube (PMT). Integration time was 0.2 seconds for each emission interval. The system was operated in the photon-counting mode.

Each emission spectrum was adjusted by subtracting the sample PMT dark current. Corion long-pass glass filters were mounted normal to the emission light path on the emission side of the sample to remove the order effects. Each emission spectrum was corrected for filter transmission to a 100 percent basis.

A sample's emission-excitation matrix (EEM) data were resampled into a 20 x 20 nm regularly-spaced XYZ data array. The new Z-coordinate value was determined using an inverse distance algorithm and the ten nearest data points to the X-Y grid coordinate. The EEM data are presented in three-dimensional (3-D) graphic plots.

A sample's potential detectivity by an airborne FLD was determined by calculating its Rhodamine WT (RhWT) dye equivalency using procedures of Hemphill, et al.⁶ The 589 nm emission of a 1200 ppb RhWT dye solution at room temperature was measured by scanning the excitation monochromator from 300-580 nm at 5 nm increments; the emission wavelength was held constant at 589 nm. This is analogous to broad band solar excitation over this spectral range. Solar weighing coefficients were calculated for 5 nm waveband intervals over the 300-850 nm solar irradiance spectrum. The weighing coefficients normalized the solar irradiance for each interval to the maximum solar irradiance at 480 nm. The solar irradiance spectrum was that of very clear atmosphere with 20 mm water vapor, 3.4 mm ozone, Angstrom turbidity coefficients of $a = 1.3$ and $b = 0.02$, and a one atmosphere light path.¹⁰ Solar weighing coefficients were applied to the RhWT dye emission spectra and the total 589 nm emission photon count was determined. The sample's EEM data were also adjusted using the solar weighing coefficients, and the RhWT dye equivalency values were calculated for the absorption lines at 423 nm, 431 nm, 486 nm, 518 nm, 589 nm, 656 nm, and 680 nm.

RESULTS AND DISCUSSION

Luminescence spectra were measured for 53 domestic and foreign camouflage fabrics, 32 paints, 39 vegetation samples, and 29 soil samples. Many of the camouflage plastic fabrics, most paints, and most soil samples had emission peaks <1000 counts, which is well below the estimated potential detection threshold of about 3200 photon counts. Most samples with emissions >1000 counts had broad band emissions in the 400-760 nm spectrum.

The samples with peak emissions >1000 counts are plotted by sample type in Figure 1: camouflage plastic fabric (G), camouflage fiber fabric (F), camouflage paint (P), vegetation (V), or soil (S). The peak emissions occurred within broad band emissions and exhibit the general similarities and differences between the sample EEMs. The excitation-emission matrices for representative samples of camouflage fabrics, paints, soil, and vegetation are illustrated in Figures 2 through 5. Each major sample category is discussed below.

Camouflage Fabrics. The two kinds of camouflage garnish fabrics had peak emissions either at Em 395-430 nm, Ex 330-370 nm; or Em 455-475, Ex 400-420 nm (Figure 1). Many plastic garnishes had peak emissions <1000 counts. These were mostly the brown-, field drab-, and olive drab-colored plastic fabrics. Some of the green-, tan-, or white-colored plastic fabrics had peak emissions >1000 counts, whereas other fabrics having these colors did not (Figure 2). Many plastic fabrics (23 of 39 samples) with emission maximums <1000 counts would not be detectable by an airborne FLD. On the other hand, many fabrics made from natural plant fibers had peak emissions >1000 counts, usually at Em 455-476, Ex 400-420 nm and would be detectable; e.g., dyed cotton fabrics used in the woodland field jacket (Figure 3).

Camouflage Paints. The chemical agent resistant coating (CARC) and alkyd paints used on military vehicles and aircraft had peak emissions in three regions; Em 370-400 nm, Ex 260-280 nm; Em 390-420 nm, Ex 330-340 nm; and Em 455-480 nm, Ex 400 nm (Figure 4). The paint samples with high emissions were the green-, tan-, or red-colored paints. Only 11 of 32 samples had emission peaks >1000 counts.

Vegetation. Emission differences between the leaf samples were associated with leaf type, leaf moisture state, and leaf pigmentation. Green-colored leaves usually had peak emissions >1000 counts, whereas the tan- and brown-colored senesced leaves had peak emissions <1000 counts. All turgid green leaves had peak emissions at Em 685 nm. The associated excitation waveband was Ex 400 nm for corn leaves, and Ex 470 nm for turgid broadleaf plants, i.e., bush bean and soybean leaflets, and pine needles. Yellow-green- or green-yellow-colored leaves had peak emissions in this same emission-excitation region, but the yellow- or brown-colored corn and soybean leaves had peak emissions at Em 475-485 nm, Ex 400 nm. Although the sampling by leaf type was not comprehensive, these data support those of Chappelle, et al.,⁷ in that graminoid leaves can be differentiated from the broadleaf and needle leaf types by the excitation-emission wavebands. The turgid green corn leaves had peak emission at Em 685 nm, Ex 400 nm, whereas the green turgid soybean and bush bean leaflets and green turgid pine needles had emission peaks at Em 685 nm, Ex 470 nm.

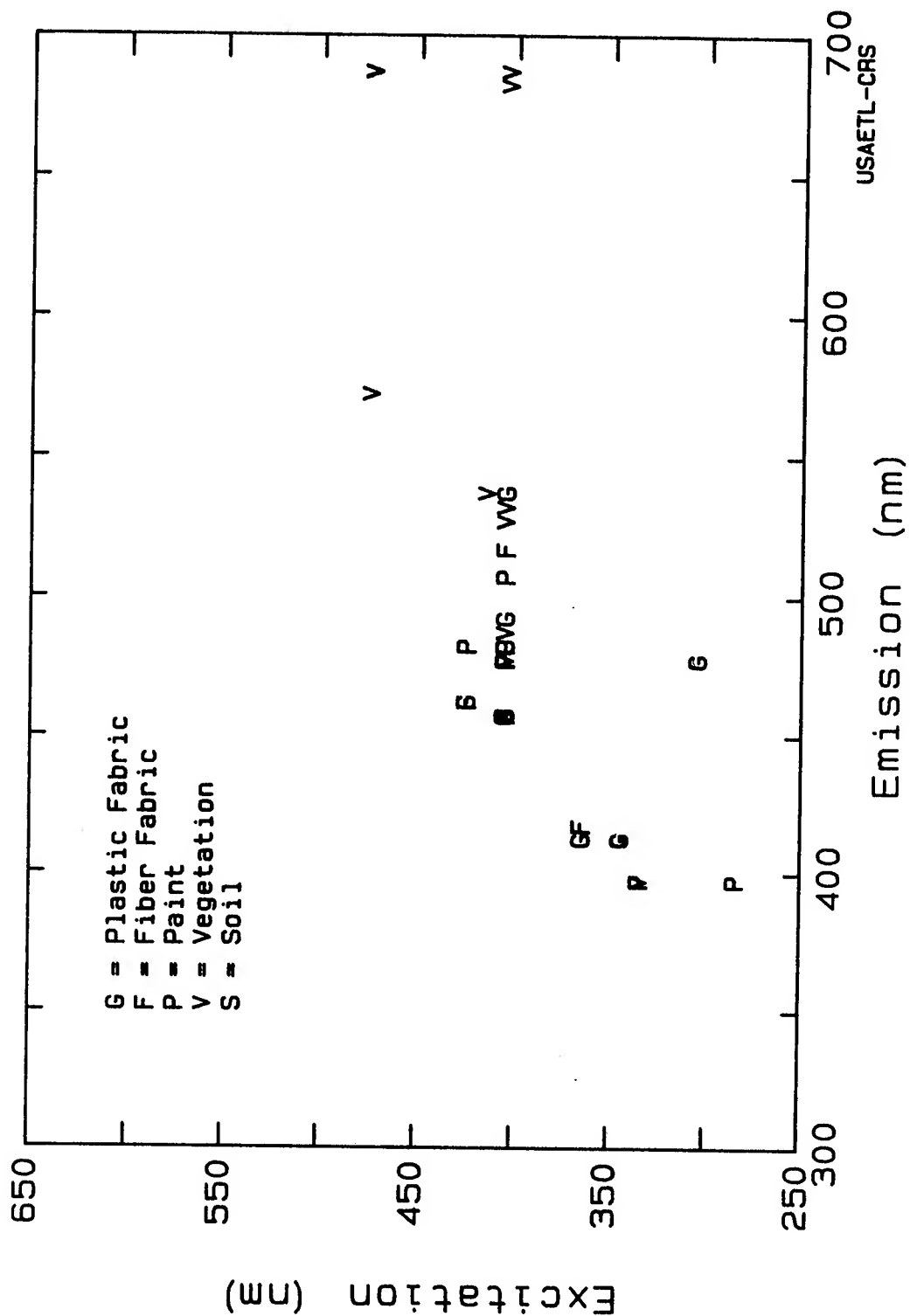
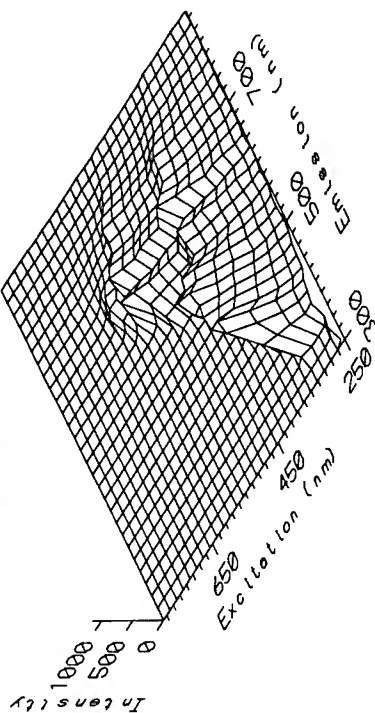


Figure 1. Camouflage Fabrics and Paints and Vegetation and Soil Samples with Emissions >1000 Counts.

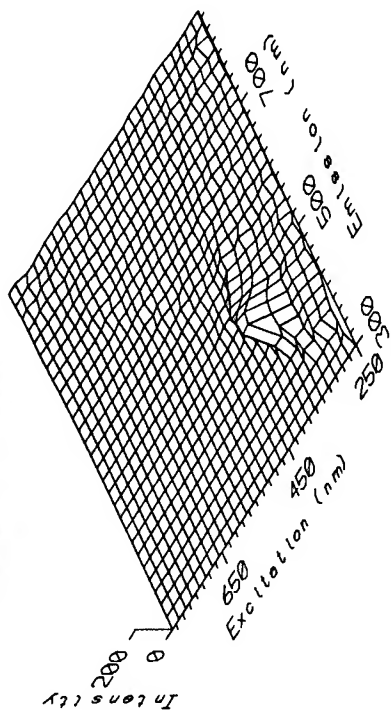
Dark Green-Colored Plastic Fabric

Emission Maximum at
Ex = 330 nm; Em = 395 nm



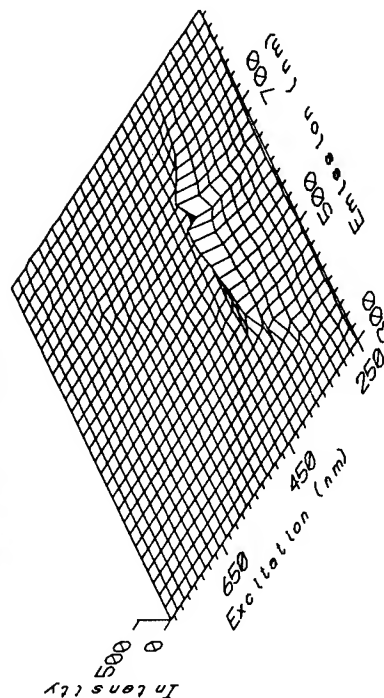
Forest Green-Colored Plastic Fabric

Emission Maximum at
Ex = 370 nm; Em = 430 nm



Tan-Colored Plastic Fabric

Emission Maximum at
Ex = 400 nm; Em = 580 nm



Brown-Colored Plastic Fabric

Emission Maximum at
Ex = 360 nm; Em = 430 nm

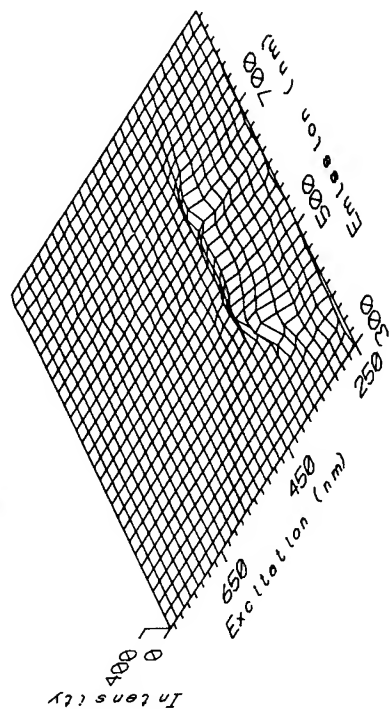


Figure 2. Excitation-Emission Matrices for Plastic Camouflage Fabrics.

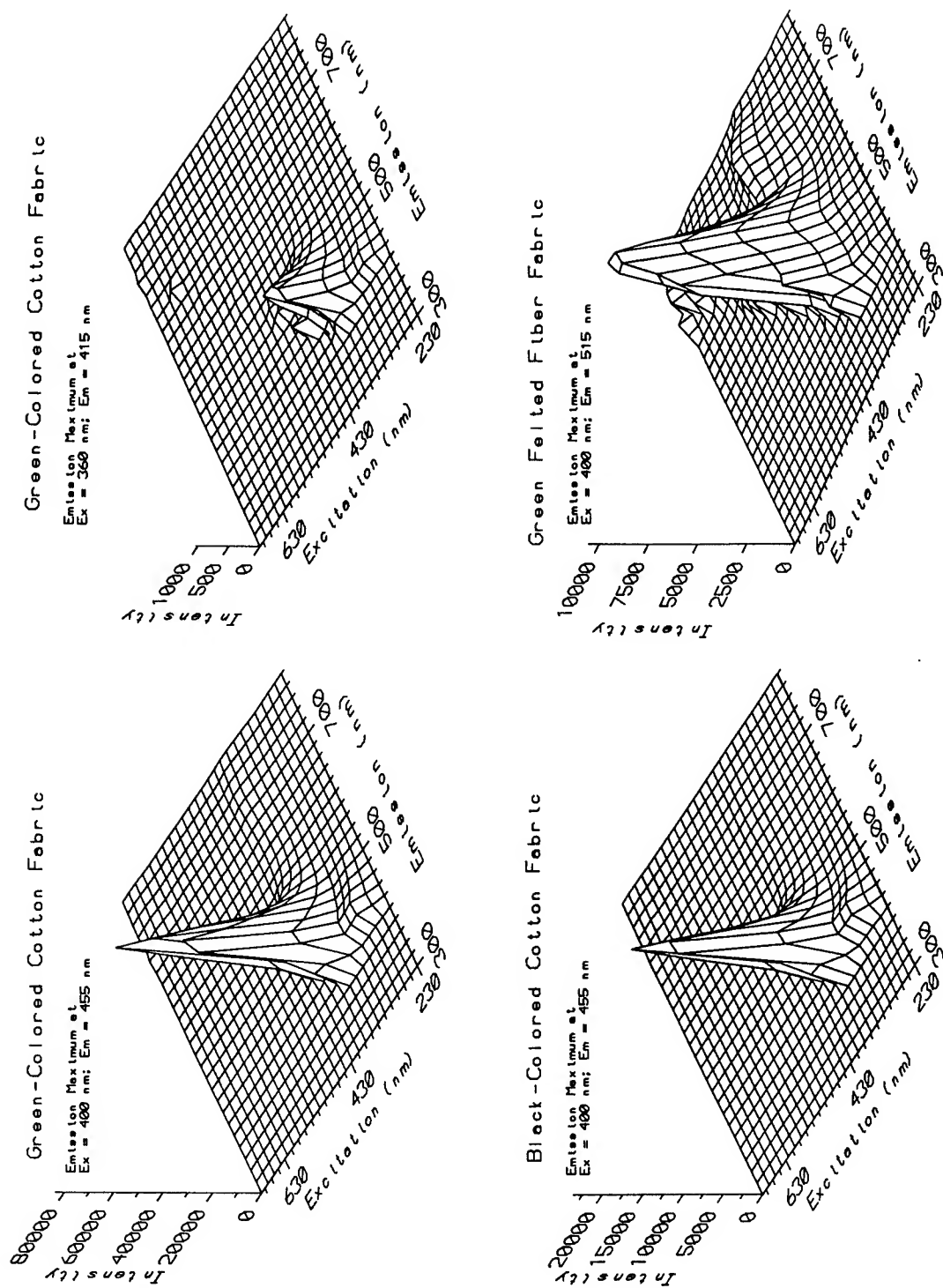
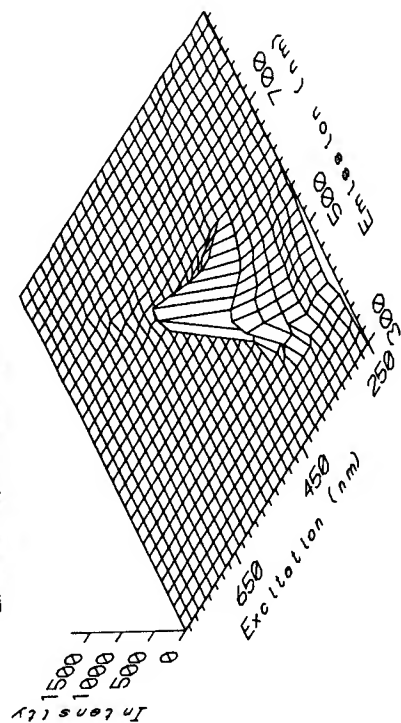


Figure 3. Excitation-Emission Matrices for Camouflage Fiber Fabrics.

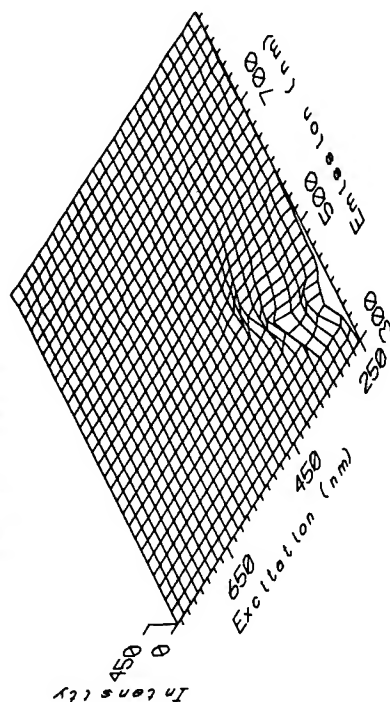
Sand-Colored CARC Paint

Emission Maximum at
Ex = 480 nm; Em = 475 nm



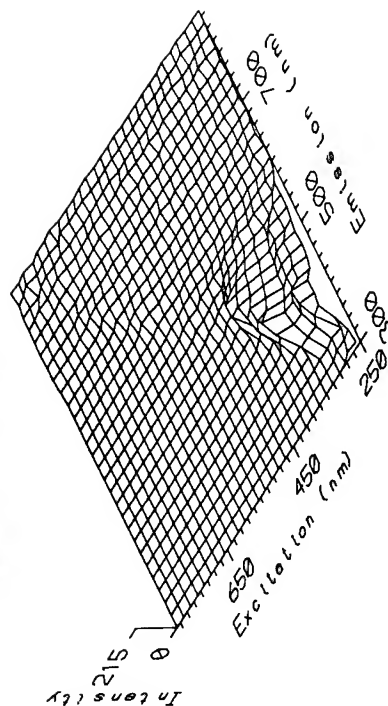
Green-Colored CARC Paint

Emission Maximum at
Ex = 340 nm; Em = 410 nm



Black-Colored CARC Paint

Emission Maximum at
Ex = 330 nm; Em = 390 nm



Brown-Colored CARC Paint

Emission Maximum at
Ex = 330 nm; Em = 445 nm

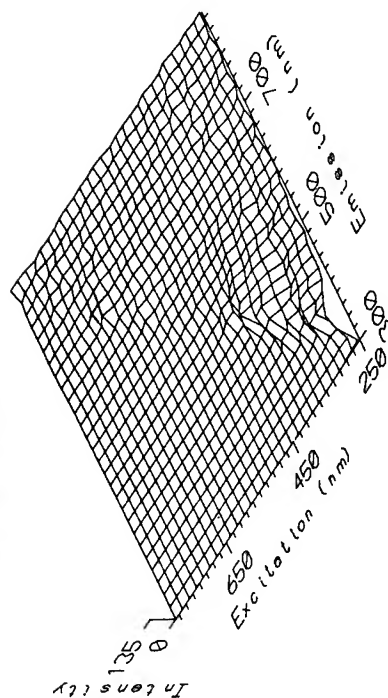


Figure 4. Excitation-Emission Matrices for Camouflage CARC Paints.

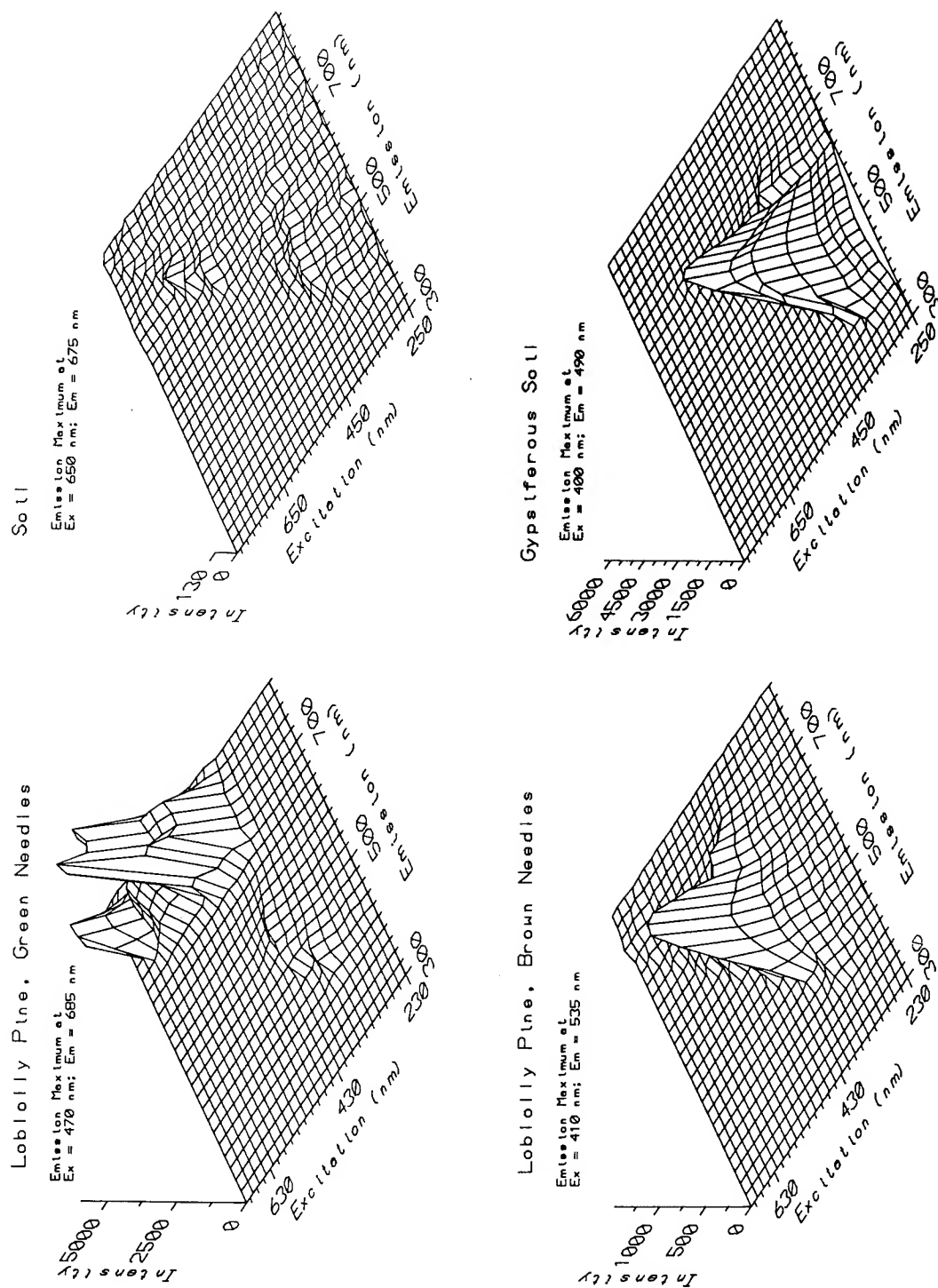


Figure 5. Excitation-Emission Matrices for Vegetation and Soil Backgrounds.

Leaf moisture affected the leaf luminescence. Upon initial drying, the emissions increased in the 685-740 nm region. Air-dried leaves had much reduced emissions in this region, but increased emissions in the 420-460 nm region. Peak emissions for the air-dried and oven-dried leaves were at 520 nm.^{2, 11}

Leaf pigmentation affected leaf luminescence as shown by the EEMs for green-, yellow-green-, green-yellow-, yellow-, and brown-colored soybean leaves and green- and brown-colored pine needles. Green- and brown-colored soybean leaves had low emission in the 460 to 540 nm region whereas the other colors of soybean leaves had emissions >1000 counts in this region. The green-yellow and yellow-colored leaflets had lower emissions in the 680-800 nm region and higher emissions in the 460-560 nm region compared with the emissions of the green-colored leaflets. The EEMs of the brown-colored soybean leaflets were very low and similar to the soil EEMs. The green- and brown-colored loblolly pine needles (Figure 5) exhibited high emissions in the 460-600 nm region, but low emissions in the 640-800 nm region.

Soils. Samples of 28 soils were measured, but 25 of these had very low emissions. The EEM of the agricultural soil shown in Figure 5 is representative of most soil samples. Unless the soil contained a high concentration of a luminescent mineral, e.g., gypsum (Figure 5), the EEM values were <700 counts.

Another method by which the sample emissions can be compared is by their computed RhWT dye equivalencies in seven absorption wavebands; 423, 431, 486, 518, 589, 656, and 680 nm. The RhWT value indicates the sample's potential detectivity by an FLD operated from an aircraft. Because the FLD is a passive-detector system, it uses the solar-stimulated luminescence that is detectable in the Fraunhofer or atmospheric absorption wavebands. In the laboratory, the FLD can detect very low concentrations of RhWT, 0.1 ppb,¹² but in practice, samples with >2.0 ppb RhWT dye equivalencies are detectable.^{9, 6}

The percentages of each sample type with RhWT equivalencies >2.0 ppm are summarized for the seven absorption wavebands in Table I. Keeping in mind that a sample's emissions are usually broad band, a sample with a detectable emission in one absorption band can also have detectable RhWT dye equivalencies in the adjacent bands. The correlations between adjacent bands are rather good, i.e., r values >0.9. The sample types were associated more with some absorption wavebands than with others. The plastic camouflage garnishes have detectable RhWT dye equivalencies in the 431, 486, and 518 nm wavebands; the fiber camouflage garnishes and camouflage paints in the 486 and 581 nm wavebands; and green-leaf vegetation in the 656 and 680 nm wavebands. The senescing and dry-leaf samples have high RhWT dye equivalencies in the <656 nm wavebands. Only the brown-colored

soybean leaf was not detectable in any absorption waveband. Two of the 14 soil samples had RhWT dye equivalencies >2.0 ppb. These were the gypsiferous soils.

Table I
Percentage of Samples with RhWT Values >2.0 ppb.

Sample Type	Total Number	Absorption Waveband (nm)						
		423	431	486	518	589	656	680
Camouflage								
Plastic	39	15	25	26	26	13	7	5
Fiber	14	36	36	50	50	43	36	43
Paint	23	4	4	30	35	26	4	-
Vegetation	39	16	18	44	44	23	87	97
Soil	14	7	7	7	7	14	7	7

The discrimination of camouflage fabrics and paints requires some emission contrast with the background materials. The EEM data graphically show that many man-made camouflage materials do not emulate either vegetation or soil backgrounds. The fabrics and paints that best emulated background surfaces were the brown- or tan-colored plastics garnishes that have low emissions and low RhWT dye equivalencies that are similar to the soil. However, some brown- or tan-colored plastic garnishes had high emissions and detectable RhWT values, making them separable from most soil backgrounds. The green-colored plastic garnishes, the natural plant fiber garnishes, and the dyed cotton fabrics did not emulate the green-colored vegetation. These garnishes luminesced in shorter wavebands than did the green leaves. The waveband emissions and RhWT dye detectivities of the green-colored fiber garnishes and camouflage paints were similar to the green-yellow- and yellow-colored leaves. The RhWT dye equivalencies show that the 486, 589, and 680 nm wavebands could be used for differentiating various camouflage fabrics and paints from vegetation and soil backgrounds.

Remote sensing of camouflage materials may require a multi-sensor system, using both luminescence and reflectance characteristics of the target and the vegetation and soil backgrounds. For example, green-colored vegetation and green-colored camouflage fabrics might be readily differentiated, even though the reflectance spectra of the camouflage nearly emulates the green-colored vegetation. These materials have high broad band emission contrasts which would permit their separation. This study has identified several wavebands in which luminescence remote sensing technology can be used to separate most camouflage materials from vegetation or soil backgrounds. Even so, there are conditions in which camouflage and the

background have very low emissions, <1000 counts, and are not separable using spectral luminescence. These conditions are identifiable and may be resolved in conjunction with reflectance remote sensing technologies. Other applications of spectral luminescence are envisioned, i.e., identifying the sample's origin, but these are limited by the availability of camouflage samples and the limited library of EEM signatures.

CONCLUSIONS

The camouflage fabrics and paints did not emulate the broad band spectral luminescence of vegetation, whereas some fabrics and paints did emulate soil. The camouflage materials have detectable emissions primarily in the 400-600 nm spectrum, which is the same region that dry or senesced leaves have detectable broad band emissions. Green-colored leaves and needles have broad band emissions in the 680-790 nm region. These emissions decreased as green-colored leaves became stressed or lost green pigmentation; this caused increased emissions in the shorter wavebands. Many camouflage paints and fabrics and the soils did not luminesce at detectable levels. In this regard, the EEMs for these samples were the same, although their colors were visually quite different.

Seven absorption wavebands were evaluated, of which the 486, 589, and 656 nm wavebands show promise for detecting and differentiating camouflage materials from background surfaces.

REFERENCES

1. Rinker, J.N., M.B. Satterwhite, J.P. Henley, and J.W. Eastes. 1989. Luminescence Measurements of Fabrics and Pigments. U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060-5546.
2. Satterwhite, M.B. 1989. Spectral Luminescence of Some Man-made Materials and Soil and Vegetation Samples. in J.N. Rinker, et al., Luminescence Measurements of Fabrics and Pigments. U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060-5546.
3. Reeves, R.G. (ed.). 1975. Manual of Remote Sensing. Amer. Soc. of Photogrammetry, Falls Church, VA., pp. 115-129.
4. Conel, J.E. and R.E. Alley. 1985. "Lisbon Valley, Utah, Uranium Test Site Report," in M.J. Abrams, J.E. Conel and H.R. Lang. The Joint NASA/GEOSAT Test Case Project, Assoc. Petrol. Geol., Vol. 1, No. 2, pp. 64-69.
5. Hemphill, W.R., G.E. Stoertz and D.A. Markle. 1969. "Remote Sensing of Luminescent Materials." Proc. 6th SRSE, pp. 656-685.

6. Hemphill, W.R., R.M. Tyson, and A.F. Theisen. 1988. "Spectral Luminescence Properties of Natural Specimens in the Scheelite-Powellite Series and an Assessment of Their Detectivity with an Airborne Fraunhofer Line Discriminator." Economic Geology, Vol. 83, pp. 637-646.
7. Chappelle, E.W. and D.L. Williams. 1987. "Laser Induced Fluorescence (LIF) for Plant Foliage." IEEE Trans. Geosci. and Remote Sensing, Vol. 25, pp. 726-736.
8. Lichtenthaler, H.K. (ed.). 1988. Applications of Chlorophyll Fluorescence. Kluwer Acad. Pub., Boston, 351 p.
9. Watson, R.D. 1981. "Quantification of Luminescence Intensity in Terms of a Rhodamine WT Standard." Lunar Planet. Inst. Tech. Rept. 81-03, pp. 19-22.
10. Mecherikunnel, A.J. and J.C. Richmond. 1980. "Spectral Distribution of Solar Radiation." NASA Tech. Memo 82021.
11. Theisen, A.F. 1988. "Fluorescence Changes of a Drying Maple Leaf Observed in Visible and Near-infrared," in H.K. Lichtenthaller (ed.), Applications of Chlorophyll Fluorescence, pp. 197-201.
12. Plascyk, J.A. and F.C. Gabriel. 1975. The Fraunhofer Line Discriminator MK-II, An Airborne Instrument for Precise and Standardized Ecological Luminescence Measurements. IEEE Trans. Instrum. Measurements IM-24, pp. 306-313.

Laser Spectroscopic and Mass Spectrometric Studies of
Propellant-Like Low Pressure Flames

*Rosario C. Sausa, Dr.

Stephen L. Howard,^a Dr.

Randy J. Locke,^a Dr.

Anthony J. Kotlar, Dr.

Andrzej W. Miziolek, Dr.

Jeffrey S. Bernstein,^b Mr.

Terrill A. Cool,^b Dr.

Ballistic Research Laboratory, SLCBR-IB-I,
Aberdeen Proving Ground, MD 21005-5066

a. NAS/NRC Postdoctoral Research Associate

b. School of Applied and Engineering Physics, Cornell
University, Ithica, NY 14853-2501

INTRODUCTION: Spatially resolved temperature profiles and profiles of major and minor species in low pressure flames are needed to better understand the chemistry occurring in flames and to validate existing chemical kinetic flame models in reference to the assumed chemical flame mechanisms as well as the accuracy of rate constant data that the models require. There has been considerable interest in recent years in the use of laser-induced fluorescence (LIF) as a technique for temperature measurements and for the monitoring of transient and stable species concentrations in flames.¹⁻⁸ It is a very attractive technique because it is nonintrusive, and offers extreme sensitivity, with detection limits down to sub-ppm levels. Also, because the laser can be focused, a spatial resolution of less than 100 microns can be achieved. The LIF technique does, however, require knowledge of the absorption and emission spectrum of the species of interest. In addition, for temperature determinations, transition probabilities and quenching rates due to the presence of other species is required, and a local thermal equilibrium must be attained so that a Maxwell-Boltzmann energy distribution analysis can be applied. Many species have been detected by LIF. The OH radical has been a particular favorite and has been the subject of the largest number of studies because it has a well known and easily accessible spectrum in the near ultraviolet, ($A^2\Sigma^+ - X^2\Pi$). Also, the OH radical is a simple diatomic which can be modeled theoretically, and is ubiquitous in flame systems.

Molecular beam sampling with triple quadrupole detection (MBS/TQD) is also a very sensitive technique that can be successfully employed for flame diagnostics. A triple quadrupole allows differentiation between species of the same mass number such as CO and N₂ or N₂O and CO₂, by mass filtering species of the same mass followed by collision induced dissociation. The fragments are then analyzed by their m/e ratio and the precursor determined. It is a very powerful technique which complements LIF since it can be employed to detect "dark" or weakly fluorescing species.

In this paper we report on the temperature profiles of low pressure stoichiometric CH₄/O₂/Ar, CH₄/NO₂, N₂O, and C₂H₄/O₂/Ar flames obtained by a Pt/Pt-Rh(10%) coated with an inert beryllium/yttrium oxide mixture thermocouple and by OH LIF rotational spectra via its (A²Σ⁺ - X²Π) (1,0) band. The spectra are fitted by a multi-variable computer program to a Maxwell-Boltzmann distribution, and a rotational temperature abstracted. These temperature profiles are then used as input for chemical kinetic flame codes from which profiles of major and transient species are generated.⁹ Species profiles are also obtained experimentally by LIF and/or molecular beam sampling with mass spectrometric detection. These profiles, in turn, are compared to those obtained from Sandia National Laboratories chemical kinetic flame models¹⁰ to test the validity of the assumed chemical flame mechanisms and accuracy of the rate constant data that the models require.

EXPERIMENTAL: The temperature and species profiles were generated in a variable pressure apparatus equipped with a flat flame McKenna burner with access for LIF, FTIR, molecular beam sampling with mass spectrometric detection, as well as thermocouple temperature measurements.¹¹ A schematic of the experimental apparatus is shown in Figure 1. Briefly, the burner with a 6-cm diameter porous stainless steel head and surrounding argon shroud is mounted to a motorized high vacuum feedthrough flange which is coupled to an off-axis rotary feedthrough. The burner can be scanned vertically with an accuracy and reproducibility of less than 50 microns and rotated with a precision of less than one degree. The burner was operated with premixed CH₄/O₂/Ar, CH₄/N₂O, NO₂, and C₂H₄/O₂/Ar stoichiometric mixtures at pressure ranging from 5-25 torr. The gas flow rates were regulated with calibrated Tylan mass flow controllers, and the pressure was regulated with a MKS pressure transducer interfaced with a MKS pressure controller and exhaust throttle valve.

Mass spectrometric detection is accomplished with a triple quadrupole manufactured by Extrel Corporation. The quartz skimmers used in the molecular beam assembly were manufactured by Florissand Scientific Glass. The orifice diameters ranged between 50 to 200 micrometers.

Platinum and platinum with 10% rhodium wires of 125 micron diameters were spot welded to produce a thermocouple junction, and mounted 180 degrees on a "V" shaped holder equipped with a spring to eliminate any sag when placed in the flame. The thermocouples were coated with a beryllium oxide(15%)/yttrium oxide mixture by a procedure reported by J.H. Kent¹² in order to avoid surface catalytic effects and to minimize embrittlement of the thermocouple resulting from high temperature contamination by a variety of elements, such as hydrogen and carbon, which occur naturally in the flame.

The tunable probe laser beam was generated by frequency doubling the output of a dye laser (Lumonics 3000) which was pumped by a Lumonics Hyper-X excimer laser operating at 308 nm with a 10 Hz repetition rate. Typical probe pulse energies at the apparatus were greater than 100 microjoules in a 5 mm diameter beam with approximately 10 nsec duration for the laser induced fluorescence detection of OH in the A-X (1,0) band near 280 nm. The laser beam line width is approximately 0.15 cm^{-1} .

The Sandia National Laboratories codes for a premixed one-dimensional laminar flame were employed to model the $\text{CH}_4/\text{O}_2/\text{Ar}$ and $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ low pressure burner supported flames. The flame model computations were performed on a VAX computer, while the spectral and temperature simulations were performed on a 386 PC-AT computer.

The laser beam was focused by a one meter lens and directed to the low pressure burner assembly containing Suprasil Brewster angle windows to minimize light scattering problems. The laser induced fluorescence was collected perpendicularly to the laser probe beam with a 300 mm lens and detected with a R-955 Hamamatsu photomultiplier tube fitted with an interference filter centered at 313 nm, 10 nm FWHM, to isolate the fluorescence from the $(v',v'') = (1,1)$ vibronic band of the OH A-X system. The output signal was fed into a boxcar averager (SRS-250) which was interfaced to a PC-AT 286 computer for data acquisition, reduction, and display. The gate was set at 9 nsec to minimize collisional quenching and/or energy transfer of the rotational excited states. In all the experiments, the laser probe energy was monitored by a Hamamatsu photodiode and was used to normalize the LIF signal. A 500 MHz digital oscilloscope was employed for rotational lifetime measurements and signal/gate display.

Argon, methane, nitrous oxide, nitrogen dioxide, and oxygen were taken from standard grade purity cylinders without purification (Matheson). The ethylene cylinder was research grade, and was obtained from Air Products.

RESULTS/DISCUSSION:

A. Temperature Profiles: The temperature profiles for $\text{CH}_4/\text{O}_2/\text{Ar}$, CH_4/NO_2 , N_2O and $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ flames at 20 torr were taken with a Pt/Pt-Rh(10%) thermocouple coated with beryllium/yttrium oxide. The ethylene containing flame is shown in Figure 2. The profile is composed of temperature measurements obtained as the burner traversed the thermocouple in the forward and reverse direction. The temperature did not vary in the two scans by more than ± 5 K. The radiation loss of the thermocouple was approximated by equating the heat transferred to the thermocouple from the gases to that lost by radiation. The corrective term is given by

$$\Delta T_{\text{corr}} = T_{\text{cal}} - T_{\text{obs}} = \epsilon \sigma d (T^4 - T_0^4) / 2h \quad (1)$$

where, ϵ is the emissivity of the coated thermocouple, σ is the Stefan-Boltzmann constant which is equal to $5.66 \times 10^{-8} \text{ W/M}^2\text{K}^4$, d the diameter of the thermocouple junction, h the heat transfer coefficient, and $T_0 = 300$ K. The diameter of the coated thermocouple was measured to be 190 microns, and an emissivity value of 0.6 was used.¹³ The heat transfer coefficient of air was employed since the correct proportions of gaseous reactants and products for a given temperature measurement is not known. Once the stable and transient species profiles are obtained a better approximation can be determined.

Also presented in Figure 2 is a temperature profile obtained by normalizing the measured thermocouple temperatures to the adiabatic temperature calculation employing Sandia's Chemical Kinetic Codes for a one dimensional pre-mixed laminar flame. The maximum measured temperature varies by only 8% from the adiabatic temperature. At this temperature the correction for radiation loss is the greatest. The variation between the two temperatures may arise from an over compensation for radiation loss due to the approximation of the emissivity and heat transfer values. A non-uniform coating will also yield a higher temperature since catalysis can occur on exposed platinum/platinum-rhodium surfaces.

OH LIF rotational spectroscopy was also employed for low pressure flame profiling. Numerous excitation scans, ranging from 281 nm - 284 nm, were taken as a function of height above the burner. A portion of the laser induced excitation spectrum of OH ($A^2\Sigma^+ - X^2\Pi$) (0,1) band near 281 nm, taken approximately 35 mm above the burner supported $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ flame is presented in Figure 3. The probe laser was attenuated to avoid saturation of the rotational transitions. The fluorescence intensity of several rotational lines on the laser energy yielded a one photon dependence. The (0,1) band was chosen over the stronger (0,0) band near 306 nm in order to minimize self absorption.

Intensity relationships can be distorted by self-absorption and a thermal equilibrium might not be established.

The OH LIF excitation scans were fitted to a Maxwell-Boltzmann rotational energy distribution using a multi-variable computer program utilizing transition probabilities and energy levels as input data. The transition probabilities were obtained from Dieke and Crosswhite.¹⁴ Some of the program variables include temperature, laser line width, line position, and laser pulse shape. The pulse shape could be approximated by either a Lorentzian or Gaussian function in either an optically thick or thin region. A portion of the fitted OH spectrum which corresponds to an excitation scan 35 mm above the 20 torr flame is also presented in Figure 3. The simulation yields a temperature of $1907 \text{ K} \pm 100 \text{ K}$, while that of the entire spectrum yields a temperature of $1952 \text{ K} \pm 100 \text{ K}$. Both temperature measurements are reported at the 95% confidence level. The close agreement in the two temperatures reveals the high degree of rotational equilibration.

The OH rotational temperature profiles for the 20 torr $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ flame are also presented in Figure 2. The profiles were generated using either a Gaussian or Lorentzian function for the approximation of the laser beam pulse shape. It can be seen that the fitted profiles agree well with the adiabatic flame temperature profile, with the maximum varying by less than 6% for the Lorentzian fit and less than 2% for the Gaussian fit.

B. Major and Transient Species Profiles: The Sandia National Laboratories flame codes have been utilized to model the $\text{CH}_4/\text{O}_2/\text{Ar}$ and $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ low pressure flames in order to better understand the detailed chemistry occurring in these flames. The mechanisms employed were that of Miller and Bowman,¹⁵ and are listed for the $\text{CH}_4/\text{O}_2/\text{Ar}$ flame in Appendix 1. The model was operated in the burner-stabilized flame mode rather than the adiabatic freely propagating flame mode. The former mode requires an experimental temperature as input into the calculation, while the latter mode requires solving the energy equation. The latter is computationally more taxing and depends in part on knowledge of the heat losses to the burner, which is difficult to calculate exactly or obtain experimentally with a high degree of accuracy.

The profiles of major and transient species, including the OH radical, generated by the flame model for the $\text{C}_2\text{H}_4/\text{O}_2/\text{Ar}$ flame at 20 torr are presented in Figure 4. The profile of the OH radical measured by LIF monitoring the $R_1(7)$ and $R_1(8)$ rotational transitions is presented in Figure 5. The $R_1(7)$ and $R_1(8)$ lines were employed since they are least sensitive to temperature variations, particularly in

regions of large temperature gradients. These rotational transitions were obtained by the Eq. (7).

$$J^2 + J - (k/hcB_v)T_{av} = 0 \quad (2)$$

where J is the rotational level, k , h , c , B_v , the usual spectroscopic constants, and T_{av} the average flame temperature. For the ethylene containing flame an average temperature of 2000 K was employed, taken from its temperature profile in Figure 2.

A comparison between the OH profile obtained experimentally to that generated by the model reveals the two contours are nearly identical, with the model even predicting the small peak occurring near the lower portion of the flame. The two profiles are, however, offset by approximately 5 mm. At present, this discrepancy cannot be explained. To further test the validity of the assumed chemical mechanisms and accuracy of the rate constants of the model more experimental species profiles are required.

Experiments aimed at generating profiles of O atoms by two photon LIF via the $2p^4 \ ^3P_2$ transition near 226 nm and CN radicals by one photon LIF via the B-X (0,0) band near 388 nm, as well as HCO and CH_3 profiles by the resonance (2+1) enhanced multiphoton ionization (REMPI) utilizing wavelengths near 397 nm and 333 nm, respectively, are now in progress. The results will be reported at the conference.

CONCLUSION: A major research facility consisting of a low pressure burner apparatus has been constructed at the Ballistic Research Laboratory for the purpose of studying the detailed flame chemistry relevant to the gaseous flames of burning propellants. Laser induced fluorescence (LIF) of OH employing its ($A^2\Sigma^+ - X^2\Pi$) (1,0) band near 281 nm is utilized for temperature measurements of low pressure $C_2H_4/O_2/Ar$, $CH_4/O_2/Ar$, and CH_4/N_2O , NO_2 flames. Temperature profiles are abstracted from a Maxwell-Boltzmann energy distribution simulation, and compared to those obtained by Pt/Pt-Rh(10%) beryllium/yttrium coated thermocouples. These temperature profiles are used as input data for the Sandia National Laboratories Chemical Kinetics Code for one dimensional premixed laminar flames to generate profiles of major and minor species. The OH radical profile has also been generated experimentally and compared to the one predicted by the model. The model is in general agreement with the experimental results. At present, however, the validity of the assumed chemical mechanisms and accuracy of the rate constants of the model cannot be fully tested because of the lack of additional species profiles. Experiments aimed at obtaining more species profiles are in progress, and will be reported at the conference.

SAUSA, HOWARD, LOCKE, KOTLAR, MIZIOLEK, BERNSTEIN & COOL

ACKNOWLEDGMENTS: We would like to thank Dr J.E.M. Goldsmith for the designs of the low pressure burner housing, and Dr. C.N Merrow for some design modifications. We also like to thank Drs. N. Witriol and W.R. Anderson for many valuable discussions. Support from BRL/NRC Postdoctoral Program (SLH & RJL), and the Army Research Office (ARO), grant number DAAI03-87-K-0066, (TAC) is acknowledged.

REFERENCES:

1. A.W. Miziolek and M.A. DeWilde, Opt. Lett., Vol. 9, p. 390, 1984.
2. B.E. Forch, J.B. Morris, and A.W. Miziolek, Laser Based Approaches on Luminescence Spectroscopy, T.Vo-Disch and D. Eastwood, Eds., ASTM Philadelphia, 1990, in press.
3. D.R. Crosely, Ed., Laser Probes for Combustion Chemistry; ACS Symposium Series; American Chemical Society, Washington, DC, Vol. 134, 1980.
4. a. D.R. Crosley and G.P. Smith, Optical Engineering, Vol. 22(5), p. 545, 1983. b. D.R. Crosely and G.P. Smith, Optical Engineering, Vol. 20, p. 511, 1981.
5. J.H. Bechtel, C.J. Dasch, and R. Teets, Combustion Research with Lasers in Laser Applications, R.K. Erf and J.F. Ready, Eds., Academic, New York, 1983.
6. A.C. Eckbreth, P.A. Bonczyk, and J.F. Verdick, Prog. Energy Comb. Sci., Vol. 5, p. 253, 1979.
7. A.C. Eckbreth, Laser Diagnostics for Combustion Temperature and Species, Abacus Press, Cambridge, MA, 1988.
8. K.J. Rensberger, J.B. Jeffries, R.A. Copeland, K. Kohse-Hoinghaus, M.L. Wise, and D.R. Crosley, Applied Optics, Vol. 28, No. 17, p. 3556, 1989.
9. To be submitted to Combustion and Flame.
10. R.J. Kee, J.F. Grcar, M.D. Smooke, and J.A. Miller, Sandia National Laboratories Report, SAND85-8240, 1985.
11. To be submitted to Review of Scientific Instruments.
12. J.H. Kent, Combustion and Flame, Vol. 14, pp. 279-282, 1970.
13. R.C. Peterson, Ph.D. Thesis, Purdue University, 1981.
14. G.H. Dieke and H.M. Crosswhite, J. Quant. Spectrosc. Radiat. Transfer, Vol. 2, p. 97, 1962.
15. Miller and Bowman, private communication.

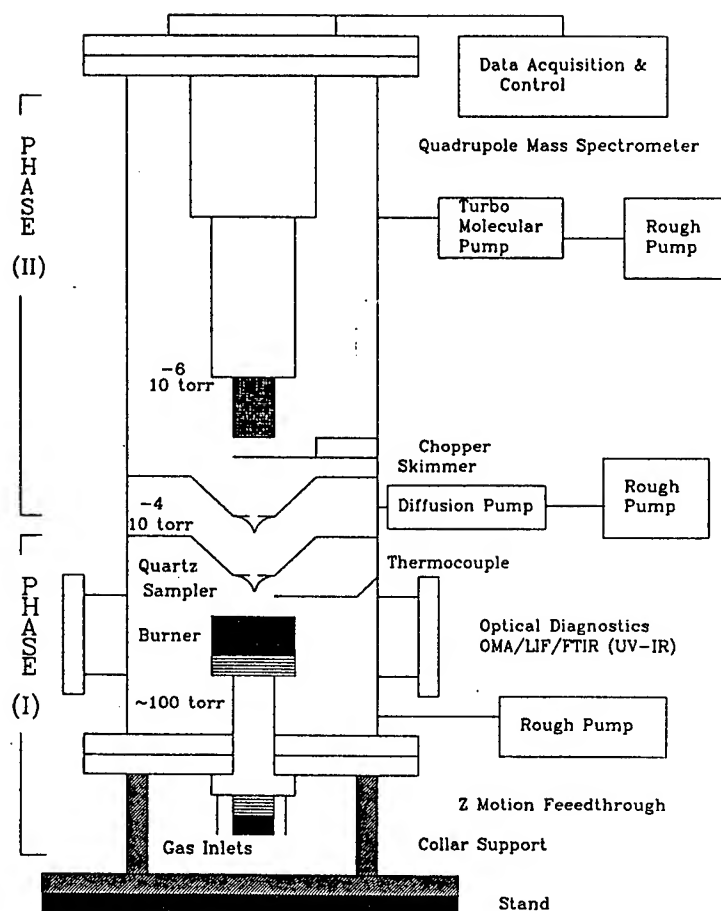


Fig. 1. Schematic of the low pressure burner system equipped for laser-based spectroscopy, FTIR, molecular beam sampling with mass spectrometric detection, and thermocouple measurements.

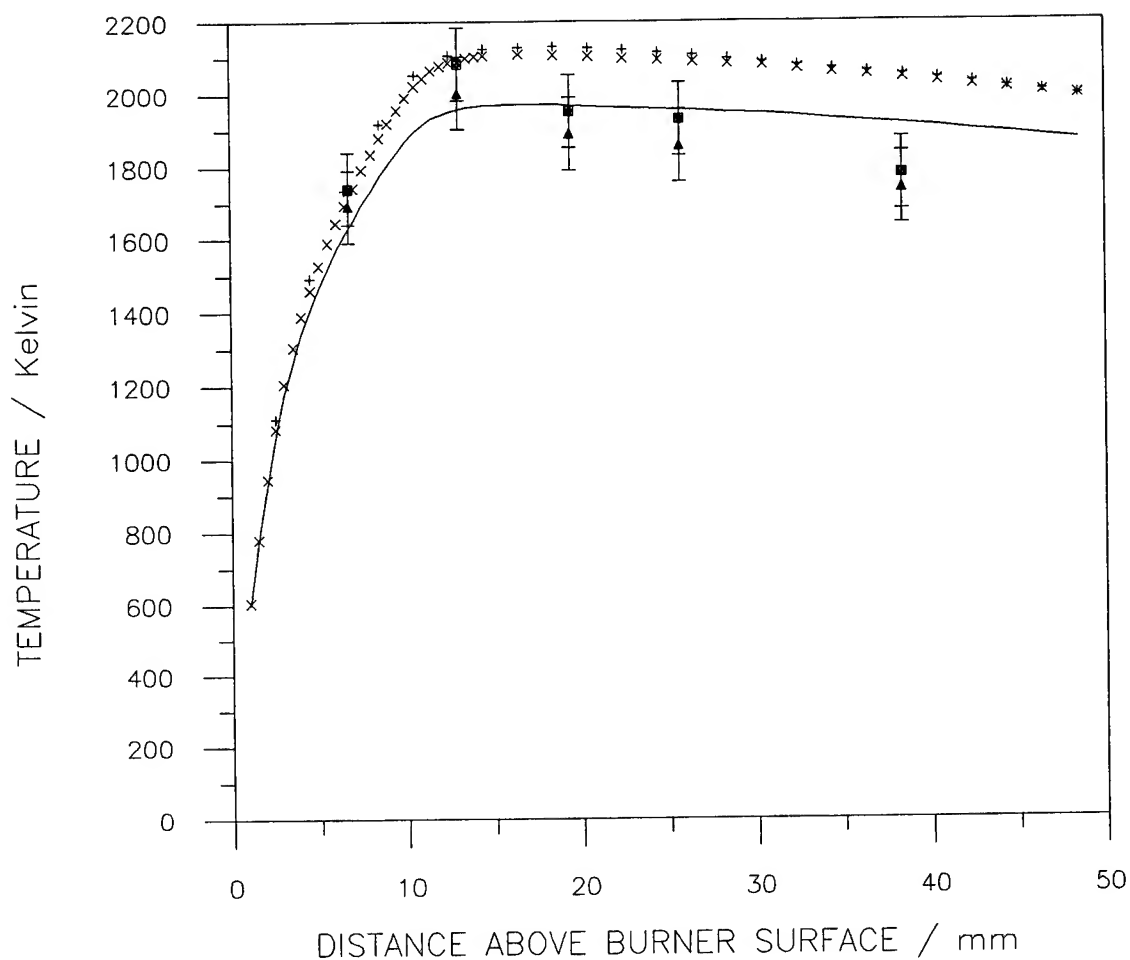


Fig. 2. Temperature profiles of a 20 torr $C_2H_4/O_2/Ar$ flame. X and + represent thermocouple temperature measurements as the burner was scanned across the thermocouple in the forward and reverse direction, respectively. The solid line is the thermocouple data normalized to a calculated value of the flame temperature, while the squares and triangles represent the temperature obtained from fitting OH LIF rotational spectra using Lorentzian and Gaussian functions, respectively, for the laser pulse shape.

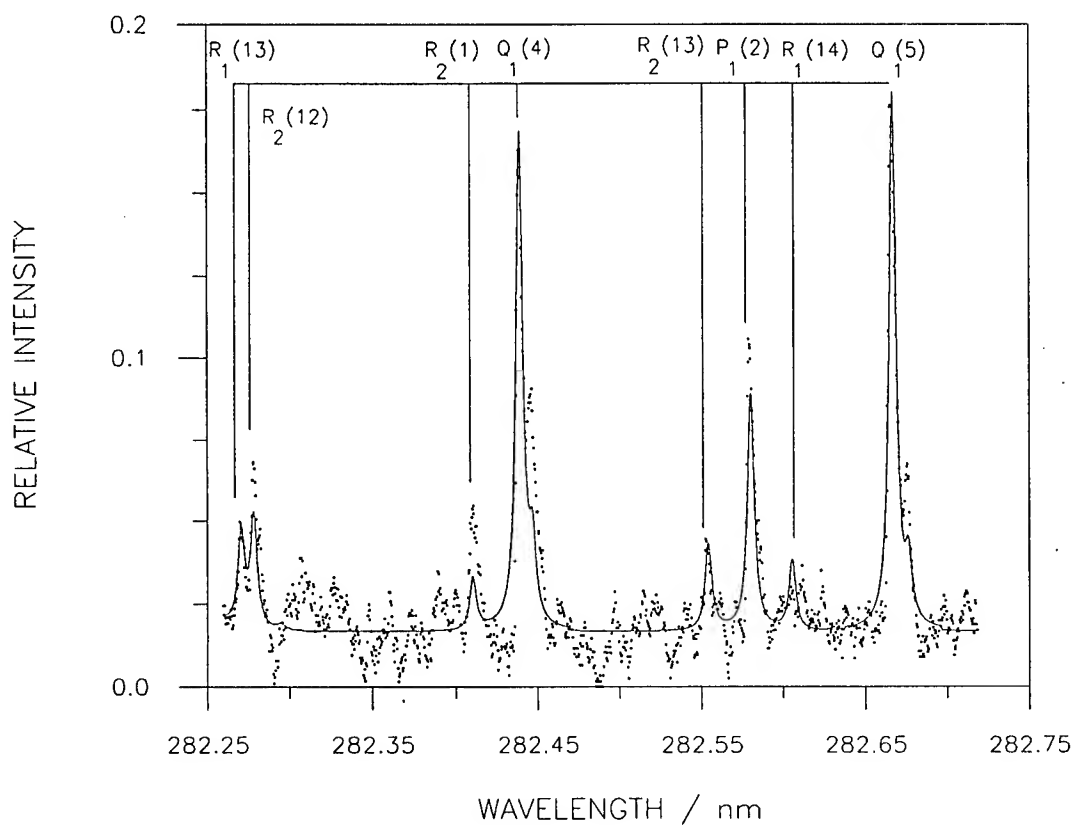


Fig. 3. A portion of the observed and simulated spectrum of the OH ($A^2\Sigma^+ - X^2\Pi$) (1,0) band. The temperature simulation of the spectrum yield $1907 \text{ K} \pm 100 \text{ K}$.

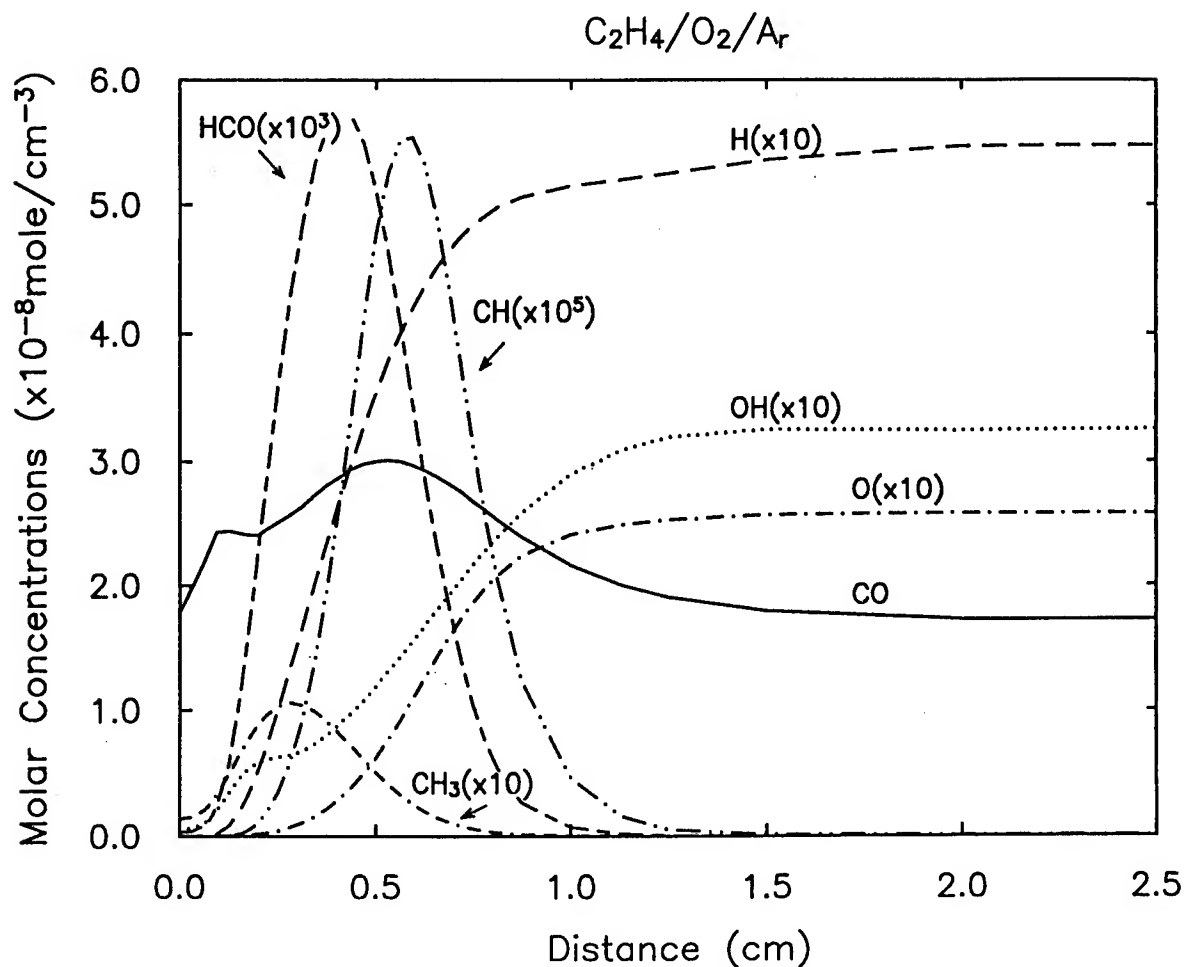


Fig. 4. Species profiles for the $C_2H_4/O_2/Ar$ 20 torr flame generated using Sandia National Laboratories chemical kinetic model for premixed one dimensional laminar flame. The mechanisms and rate constants employed were obtained from Miller and Bowman.

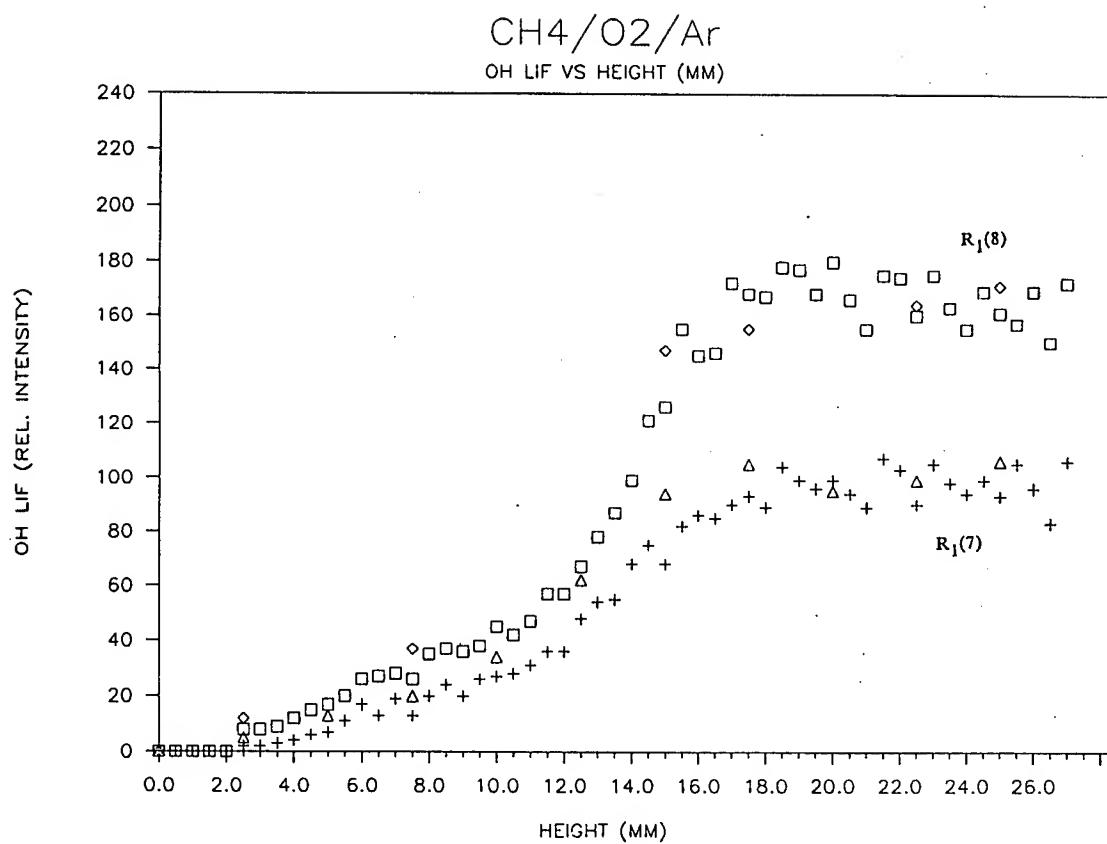


Fig. 5. OH LIF profile obtained by monitoring the $R_1(7)$ and $R_1(8)$ rotational transitions of the $A^2\Sigma^+ - X^2\Pi$ (1,0) band near 281 nm.

APPENDIX 1. Table of modified reactions for the CH₄/O₂/Ar 20 torr flame employed in the chemical kinetic flame model.

```

!!!!!! MODIFIED REACTIONS FOR CH4/O2/AR OF MILLER(1988)!!!!!!
ELEMENTS H O C AR END
SPECIES H2 O2 H2O C2H4 CO2 CO CH3 AR H O HO2 OH CH2 CH H2O2 CH2O HCO
      C CH4 C3H2 CH3O C2H C2H2 C2H3 C2H5 C2H6 CH2CO
END
REACTIONS
!
CH4+O2=CH3+HO2      7.9E13      0.0      56000.
CH4+H=CH3+H2        2.2E4       3.0      8750.
CH4+O=CH3+OH         1.6E6      2.36     7400.
CH4+OH=CH3+H2O       1.6E6      2.1      2460.  !!
!
CH3+H(+M)=CH4(+M)    6.0E16     -1.0     0.0  !!!!! FALL-OFF
      LOW/ 8.0E26      -3.0      0.0 /
      H2/2.0/ CO/2.0/ CO2/3.0/ H2O/5.0/
CH3+O=CH2O+H         6.8E13      0.0      0.0  !!!!!
CH3+O=CH2+OH          5.0E13      0.0     12000.  !!!!!
CH3+OH=CH2O+H2        1.0E12      0.0      0.0
CH3+OH=CH2+H2O         1.5E13      0.0     5000.
CH3+H=CH2+H2          9.0E13      0.0     15100.
CH3+M=CH2+H+M         1.9E16      0.0     91600.
CH3+O2=CH2O+OH         5.2E13      0.0     34570.  !
CH3+O2=CH3O+O          7.0E12      0.0     25650.  !
!
CH2+H=CH+H2          1.4E19      -2.0      0.0  !!!!!
CH2+OH=CH2O+H         2.5E13      0.0      0.0  !!!!!
CH2+OH=CH+H2O          4.5E13      0.0     3000.  !!!!!
CH2+CH=C2H2+H          4.0E13      0.0      0.0
CH2+CH2=C2H2+H2        3.2E13      0.0      0.0
!
CH+O2=HCO+O           3.3E13      0.0      0.0
CH+O=CO+H             5.7E13      0.0      0.0
CH+OH=HCO+H           3.0E13      0.0      0.0
CH+CO2=HCO+CO         3.4E12      0.0     690.
CH+CH3=C2H3+H          3.0E13      0.0      0.0
CH+CH4=C2H4+H          6.0E13      0.0      0.0
CH+C2H2=C3H2+H         1.3E14      0.0      0.0  !!
!
CH2+CO2=CH2O+CO        1.1E11      0.0     1000.
CH2+O=CO+H+H          3.0E13      0.0      0.0
CH2+O=CO+H2           5.0E13      0.0      0.0
CH2+O=CH+OH           5.0E13      0.0     12000.
CH2+O2=CO2+H+H        1.6E12      0.0     1000.
CH2+O2=CH2O+O          5.0E13      0.0     9000.
CH2+O2=CO2+H2         6.9E11      0.0      500.
CH2+O2=CO+H2O          1.9E10      0.0    -1000.
CH2+O2=CO+OH+H         8.6E10      0.0     -500.
CH2+O2=HCO+OH          4.3E10      0.0     -500.
!
CH3O+M=CH2O+H+M        1.0E14      0.0     25000.  !
CH3O+H=CH2O+H2         2.0E13      0.0      0.0  !
CH3O+OH=CH2O+H2O       1.0E13      0.0      0.0  !
CH3O+O2=CH2O+HO2       6.3E10      0.0     2600.  !
!
CH2O+OH=HCO+H2O        3.43E9      1.18    -447.
CH2O+H=HCO+H2          2.19E8      1.77    3000.  !
CH2O+M=HCO+H+M         3.31E16      0.0     81000.
CH2O+O=HCO+OH          1.81E13      0.0     3082.
!
HCO+OH=CO+H2O          5.0E12      0.0      0.0
HCO+M=H+CO+M           2.5E14      0.0     14700.0  !!!!!
      CO/1.9/ H2/1.9/ CH4/2.8/ CO2/3.0/ H2O/5.0/
HCO+H=CO+H2           4.0E13      0.0      0.0
HCO+O=CO2+H            1.0E13      0.0      0.0  !
HCO+O=CO+OH            3.0E13      0.0      0.0  !
HCO+O2=HO2+CO          3.3E13     -0.4      0.0
!
C+OH=CO+H             5.0E13      0.0      0.0  !
C+O2=CO+O              2.0E13      0.0      0.0  !
C+CO2=2CO               6.0E8      0.0      0.0  !
C+CH4=CH+CH3           5.0E13      0.0     24000.  !
C+CH3=C2H2+H           5.0E13      0.0      0.0  !! C-2
C+CH2=C2H+H            5.0E13      0.0      0.0  !! C-2

```

```

!
CO+O+M=CO2+M      3.2E13      0.0      -4200.
CO+OH=CO2+H        1.51E7      1.3      -758.
CO+O2=CO2+O        1.6E13      0.0      41000.  !!
CO+HO2=CO2+OH       5.8E13      0.0      22934.
H+H+CO2=H2+CO2     5.49E20     -2.0      0.0
!!!! C-2 MECH *****
CH3+CH3(+M)=C2H6(+M) 9.3E16     -1.2     654.  !!!! FALL-OFF
LOW / 3.18E41      -7.0     2762 /
H2/2.0/ CO/2.0/ CO2/3.0/ H2O/5.0/
CH3+CH3=C2H4+H2     2.1E14      0.0     19200.
CH3+CH2=C2H4+H       3.0E13      0.0      0.0
!
C2H6+H=C2H5+H2       5.4E2      3.5     5200.
C2H6+O=C2H5+OH       2.5E13      0.0     6360.
C2H6+OH=C2H5+H2O     8.7E9       1.05    1810
C2H6+CH3=C2H5+CH4    5.5E-1       4.0     8280.
C2H6+CH2=CH3+C2H5    2.2E13      0.0     8660.
!
C2H5+H=CH3+CH3       4.0E13      0.0      0.0
C2H5+M=C2H4+H+M      1.0E17      0.0     31000.  !! C2H4+H=C2H5, 1.0e13 0.0 1500.
C2H5+O2=C2H4+HO2     3.2E12      0.0     5020.
!
C2H4+M=C2H2+H2+M     2.6E17      0.0     79350.
C2H4+M=C2H3+H+M      2.6E17      0.0     96600.
C2H4+H=C2H3+H2       1.1E13      0.0     10000.  !!! #74 high sensitivity E14
C2H4+O=HCO+CH3       1.6E9       1.2      746
C2H4+OH(+M)=C2H3+H2O(+M) 2.0E13      0.0     5955  !!!!! FALL-OFF/WD(1988)
LOW / 2.9E11      0.0     1570. /
C2H4+OH=CH2O+CH3     2.0E11      0.0      960.  !!!!! #77 high sensitivity 2.0E12
!
C2H3+M=C2H2+H+M      8.0E14      0.0     31500.
C2H3+H=C2H2+H2       4.0E13      0.0      0.0
C2H3+O=CH2CO+H       3.3E13      0.0      0.0
C2H3+OH=C2H2+H2O     5.0E12      0.0      0.0
C2H3+O2=HCO+CH2O     4.0E12      0.0     -250.  !!!! NEW GUTMAN(1984)
C2H3+CH2=C2H2+CH3    3.0E13      0.0      0.0
C2H3+C2H=C2H2+C2H2   3.0E13      0.0      0.0
!
C2H2+M=C2H+M+H       4.2E16      0.0     107000.
C2H2+O=CH2+CO        2.2E10      1.0     2580.
! C2H2+O=HC2O+H       3.6E4       2.7     1390.
C2H2+O=C2H+OH        3.2E15      -0.6     15000.
C2H2+OH=CH2CO+H      3.2E11      0.0      200.
C2H2+OH=C2H+H2O      6.0E12      0.0     7000.
!
C2H+H2=C2H2+H        4.1E5       2.39     860.
C2H+O=CH+CO          5.0E13      0.0      0.0
! C2H+OH=HC2O+H      2.0E13      0.0      0.0
C2H+O2=CO+HCO        2.4E12      0.0      0.0
!C2H+O2=HC2O+O       6.0E11      0.0      0.0
!
CH2CO+M=CH2+CO+M     3.6E15      0.0     59300.
CH2CO+H=CH3+CO       1.1E13      0.0     3430.
! CH2CO+H=HC2O+H2     7.5E13      0.0     8000.
CH2CO+O=CH2O+CO      2.0E13      0.0      0.0
! CH2CO+O=HC2O+OH     5.0E13      0.0     8000.
CH2CO+OH=CH2O+HCO    2.8E13      0.0      0.0
!CH2CO+OH=HC2O+H2O   7.5E12      0.0     3000.
!
!HC2O+H=CH2+CO       1.1E14      0.0      0.0
!HC2O+O=2CO+H        1.1E14      0.0      0.0
!HC2O+OH=HCO+CO+H    1.0E13      0.0      0.0
!HC2O+O2=2CO+OH      1.5E12      0.0     2500.
!HC2O+CH2=C2H+CH2O   1.0E13      0.0     2000.
!HC2O+CH2=C2H3+CO    3.0E13      0.0      0.0
!HC2O+HC2O=C2H2+2CO  1.0E13      0.0      0.0
!
!!!! H2/O2 FLAME Ref. Miller Chemkin Sample(1988)////
H2+O2=2OH            1.7E13      0.0     47780.
OH+H2=H2O+H          1.17E9      1.3     3626.  !D-LSW
H+O2=OH+O             5.13E16     -0.816   16507.  !!!!! JAM,JCP 1981
O+H2=OH+H            1.8E10      1.0     8826.
H+O2+M=HO2+M         3.61E17     -0.72     0.  !DIXON-LEWIS

```

SAUSA, HOWARD, LOCKE, KOTLAR, MIZIOLEK, BERNSTEIN & COOL

H2O/18.6/	H2/2.86/	CO2/4.2/	CO/2.11/	O2/0.0/	AR/1.26/
H+O2+O2=HO2+O2	6.7E19	-1.42	0.	!SLACK, JAN	
OH+HO2=H2O+O2	7.5E12	0.0	0.		
H+HO2=2OH				1.4E14	0.0 1073.
O+HO2=O2+OH				1.4E13	0.0 1073.
2OH=O+H2O	6.0E8	1.3	0.	!COHEN-WEST.	
H+H+H2=H2+H2	9.2E16	-0.6	0.0		
H+H+H2O=H2+H2O	6.0E19	-1.25	0.0		
O2+M=O+O+M		1.85E11	0.5	95560.	! (?)
H+O+M=OH+M			6.2E16	-0.6	0.0
H2O/5/					
H+OH+M=H2O+M			1.6E22	-2.0	0.
H2O/5/					
H+HO2=H2+O2			1.25E13	0.0	0.
HO2+HO2=H2O2+O2			2.0E12	0.0	0.
H2O2+M=OH+OH+M			1.3E17	0.0	45500.
H2O2+H=HO2+H2			1.6E12	0.0	3800.
H2O2+OH=H2O+HO2			1.0E13	0.0	1800.
END					

Excimer Laser Photodissociation of Selected Oximes:
Search for H_2CN Radical

Rosario C. Sausa, Dr.

*Andrzej W. Miziolek, Dr.

William R. Anderson, Dr.

Paul J. Dagdigian,^a Dr.

Ballistic Research Laboratory, SLCBR-IB-I,

Aberdeen Proving Ground, MD 21005-5066

a. The Johns Hopkins University, Baltimore, MD 21218

I. INTRODUCTION: There has been considerable interest in the development of sensitive laser spectroscopic probes for the detection of trace transient species in combustion environments.^{1,2} Laser-induced fluorescence has been utilized for the observation of a number of free radical combustion species.³ Methyleneamidogen was first observed by Cochran, et al.,⁴ by ESR spectroscopy in an argon matrix. The ultraviolet absorption spectrum of H_2CN was first observed through flash photolysis of formaldazine [$(\text{H}_2\text{CN})_2$]⁵⁻⁷ and subsequently in the photolysis of formaldoxime.⁸ Methyleneamidogen is believed to be formed in the early stages of the decomposition of nitramines, a class of important high-energy molecules.⁹⁻¹⁰ Using ESR spectroscopy, Morgan and Beyer¹¹ observed H_2CN , along with nitrogen dioxide, as one of the species present in the vapors produced by the slow pyrolysis of cyclotetramethylenetetranitramine (HMX) near its melting point. This transient molecule is also believed to be formed by the dissociation of methylenenitramine [CH_2NNO_2], which has been experimentally identified as a primary decomposition product in the molecular beam infrared multiphoton dissociation of cyclotrimethylenetri-nitramine (RDX).¹²

In this paper, we report on our attempt to observe laser-induced fluorescence of the methyleneamidogen radical, H_2CN , which is prepared by excimer laser photodissociation of formaldoxime [H_2CNOH]. We have also investigated the photodissociation of the methylated homologs of formaldoxime, namely acetaldoxime [CH_3CHNOH] and acetoxime [$(\text{CH}_3)_2\text{CNOH}$]. We also report the internal state distributions of the companion photofragment, hydroxyl, for these three precursors. Detection of the hydroxyl radical allows indirect verification of the presence of the oxime precursor and also provide a way to estimate our detection sensitivity for methyleneamidogen.

II. EXPERIMENTAL SECTION: These experiments were carried out in a large vacuum chamber normally used for molecular beam scattering studies.¹³ The photolysis source was an excimer laser (Lambda Physik EMG101MSC) usually operated at 193 nm with ArF; typical pulse energies of the unpolarized output were 10-20 mJ in a 1 cm x 3 cm rectangular area at the photolysis zone. A few experiments with acetaldehyde precursor were carried out at 248 nm with KrF. The tunable ultraviolet probe laser beam, obtained by frequency doubling the output of a Nd:YAG pumped dye laser (Quintel), crossed at right angles to the excimer laser beam along the long dimension. Typical probe pulse energies at the apparatus were 200 μ J in a 4 mm diameter beam for laser fluorescence detection of OH in the $A^2\Sigma^+ - X^2\Pi$ (1,0) and (2,1) bands at 280-290 nm. Slightly higher power (500-800 μ J) was employed in the search for fluorescence from H_2CN and its homologs. Fluorescence detection of CN through excitation of its $B^2\Sigma^+ - X^2\Sigma^+$ (0,0) band near 388 nm was accomplished by mixing the dye laser output with the residual 1.06 μ m Nd:YAG fundamental. The incident pulse energy of the probe radiation was approximately 100 μ J.

Fluorescence excited by the probe laser was collected with a 3-lens telescope and was detected with a photomultiplier (EMI 9813QB), whose output was directed to a gated integrator (Stanford Research Systems). For experimental convenience, the fluorescence telescope has a 90° bend in it. To reduce the scattered light background, a dichroic mirror, with reflectivity peaked over 300-350 nm, was used to make this bend. Below 300 nm, the reflectivity dropped rapidly, reaching 50% at 285 nm. For fluorescence detection of CN, the dichroic mirror was replaced with an aluminized one, and excimer laser scattered light was eliminated with the insertion of a 390 nm center wavelength, 10 nm FWHM filter. In many runs, excitation spectra were acquired under computer control (DEC LSIP-11/23), and the spectra were stored on magnetic diskettes for later analysis on another laboratory computer (Apple Macintosh).

The precursors acetaldoxime and acetoxime were obtained from Aldrich Chemicals. The stated purities were 95 and 98%, respectively. Formaldoxime was synthesized by the procedure of Scholl.¹⁴ The identity and purity of the solid product was checked by infrared and gas chromatograph/mass spectra. Acetaldoxime and acetoxime, which are monomeric precursors, were admitted into the vacuum chamber by means of a needle valve on an evacuated flask containing the degassed material. Formaldoxime exists as a solid polymer at normal conditions (298 K and 1 atm), and was heated to 80-90°C to produce a pressure of several mTorr in the chamber with the pumps on.

III. RESULTS AND DISCUSSION:

A. Fragmentation Energetics: There is considerable uncertainty in the $R_1R_2\text{CN-OH}$ bond dissociation energy for the oximes. The bond dissociation energy for formaldoxime ($R_1 = R_2 = \text{H}$) can be estimated by using its heat of formation ($\Delta H_{f298}^0 = 0$ kcal/mole) derived from bond additivity considerations by Benson and O'Neal¹⁵ and the H-HCN binding energy (19 kcal/mole) calculated by Bair and Dunning¹⁶ for the H_2CN species. With the aid of the well-determined heats of formation of H and HCN ,¹⁷ we estimate a bond energy of 74 kcal/mole for formaldoxime. The Bond-Additivity-Corrected Moller-Plesset fourth order perturbation theory (BAC-MP4) method has been used to calculate the thermochemistry of a number of combustion species.¹⁰ Using the BAC-MP4 values¹⁸ for ΔH_{f298}^0 of H_2CNOH and H_2CN (5.1 and 58.9 kcal/mole, respectively), we obtain a $\text{H}_2\text{CN-OH}$ bond energy of 63.0 kcal/mole, which is in moderate agreement with the previous cruder estimate.

We would expect methylation not to change drastically the $R_1R_2\text{CN-OH}$ bond energy. Benson and O'Neal¹⁵ estimated a value of 48.4 kcal/mole for acetaldoxime ($R_1 = \text{CH}_3$, $R_2 = \text{H}$) from available experimental unimolecular decay kinetic data. This value seems quite low in light of the estimates for formaldoxime. Indeed, Benson and O'Neal questioned the validity of the experimental data. We can obtain another value using their estimated¹⁵ ΔH_{f298}^0 for acetaldoxime (-7.3 kcal/mole) and a BAC-MP4 value¹⁸ for CH_3CHN (50.1 kcal/mole). We obtain a $\text{CH}_3\text{CHN-OH}$ bond energy of 66.6 kcal/mole, which is quite similar to that for formaldoxime. At present, there is no information on acetoxime ($R_1 = R_2 = \text{CH}_3$); however, we expect a similar bond energy.

The ultraviolet spectra of the oximes contain two absorption regions, a weak diffuse band near 210-213 nm and an intense band with a maximum below 190 nm.⁶ Accordingly, our photodissociation studies were carried out only with an ArF excimer laser (193 nm) as the photolysis light source. This implies that there is approximately 85 kcal/mole of energy available to the photofragments, if a $R_1R_2\text{CN-OH}$ bond energy of 63 kcal/mole is assumed.

B. OH Fragment Internal State Distributions: The OH product was observed by laser fluorescence excitation in the $A^2\Sigma^+ - X^2\Pi$ (1,0) and (2,1) bands. These particular bands were chosen since they lie very close to the wavelengths of the absorptions previously observed for H_2CN and its homologs. The dependence of the OH fluorescence signals for several lines was investigated as a function of excimer laser pulse energy. On a log-log plot the slope was found to equal 1.25 ± 0.1 . This is roughly consistent with a one-photon photodissociation process. Under thermalized conditions, the OH fluorescence signal decreased with a half life of approximately 100 μsec with respect to the

pump/probe delay in 0.7 Torr nitrogen buffer gas. This is comparable with the expected diffusion time out of the photolysis zone.

The OH($v=0$) rotational distributions were taken from (1,0) R_1 and R_2 branch intensities. The intensities were converted to rotational populations using fluorescence excitation line strength factors calculated by the formulas of Greene and Zare.¹⁹ It was assumed that the OH photofragments had an isotropic M_J distribution. The nascent distributions, which are plotted in Figure 1, were found to be somewhat hotter than for thermalized samples. The distributions are essentially identical for all three oxime precursors and exhibit a preference for production of the $F_1(\Omega = 3/2)$ over the $F_2(\Omega = 1/2)$ spin-orbit component, particularly at low N .

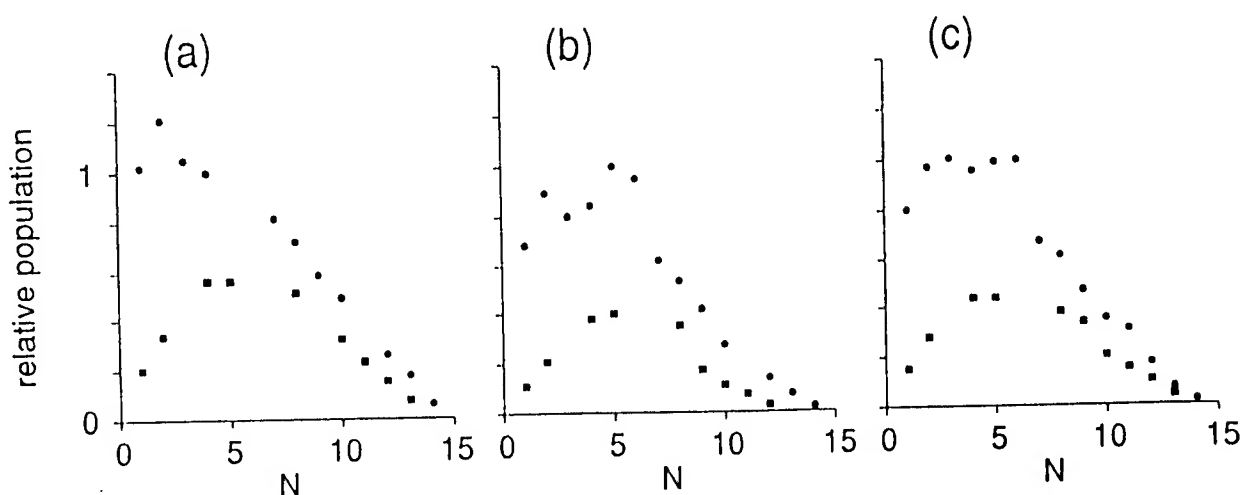


Figure 1. Rotational State Populations for OH ($v=0$) Produced in the 193-nm Photodissociation of (a) Formaldoxime, (b) Acetaldoxime, and (c) Acetoxime. The circles and squares represent the $F_1(\Omega=3/2)$ and $F_2(\Omega=1/2)$ spin-orbit levels, respectively. The distributions are each separately normalized to a particular F_1 level. The populations are not divided by the $(2J+1)$ degeneracy factor.

The OH($v=1$)/OH($v=0$) population ratio was also estimated for acetaldoxime precursor since high- N P_1 and P_2 lines of the (1,0) band overlapped the (2,1) band origin, as shown in Figure 2. We estimate that the nascent $v = 1$ to $v = 0$ ratio equals 0.02 ± 0.01 . In deriving this ratio, a correction for predissociation^{20,21} of the low N levels of $v'=2$ was made. From the measured OH internal state distribution, we conclude that an average of only 2.5 kcal/ mole appears as excitation of the OH fragments for acetaldoxime precursor. Similar results apply to

formaldoxime and acetoxime. Thus, the overwhelming majority of the energy available to the fragments must be present as translational recoil and internal excitation of the R_1R_2CN fragment.

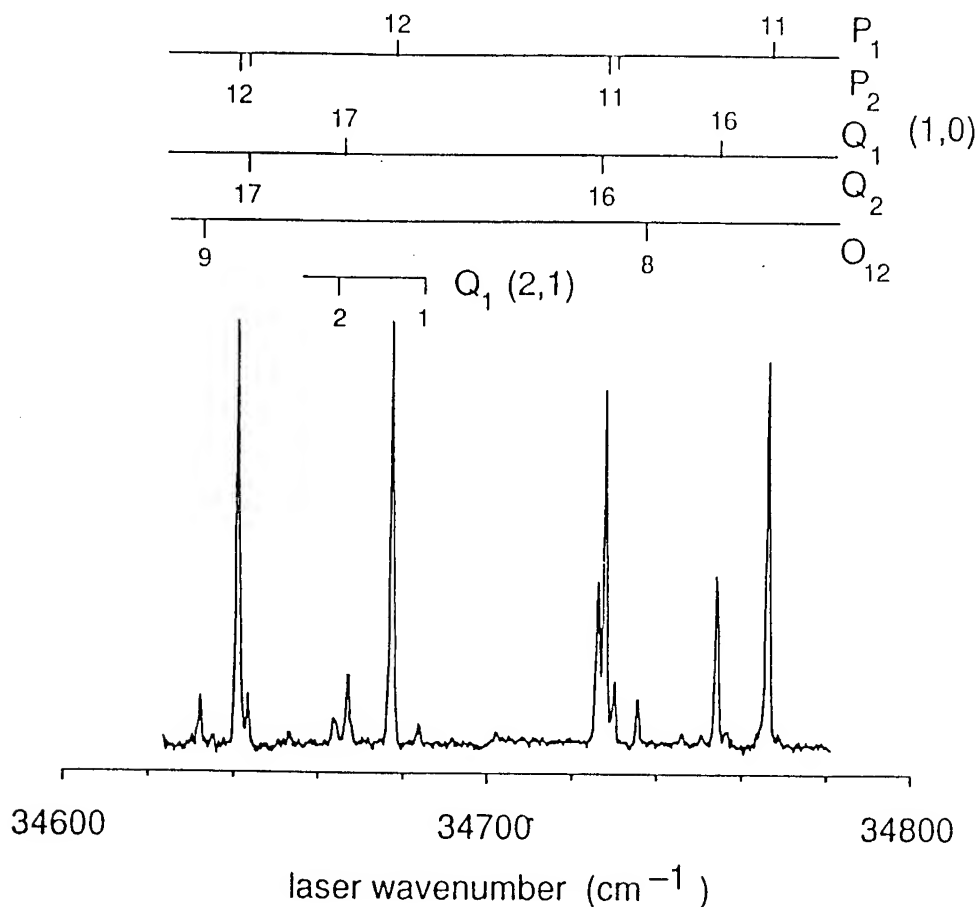


Figure 2. Excitation Spectrum Near the OH A-X (2,1) Band Origin for Photodissociation of Acetaldoxime Under Nascent Conditions. Lines of the (1,0) band and of the (2,1) Q₁ branch are marked. The pressure and pump/probe delay were 4 mTorr and 3 μ s, respectively.

C. Search for H₂CN Fluorescence: With OH detected as a photofragment from the oximes, it can be assumed that H₂CN, or its homologs, will be present in the photolysis at the same concentration as OH if no subsequent excimer-laser-induced decomposition or reactions of R_1R_2CN occur. (In Section D, we address the possibility that this

assumption may not be correct.) Extensive searches for laser fluorescence signals from H_2CN and its homologs were carried out, both under nascent and thermalized conditions. No fluorescence signals attributable to these species could be found. For the formaldoxime precursor, scans were taken over 280-286 nm in the region of the strongest absorption bands seen in the flash photolysis experiments on formaldoxime.^{5,6,8} Because of the ease of introducing the precursor into the apparatus, the photolysis of acetaldoxime and acetoxime were more extensively studied. For these molecules, scans were taken over the range 287-295 nm in the regions where CH_3CHN and $(\text{CH}_3)_2\text{CN}$ have been found to absorb.⁷

From our lack of observation of $\text{R}_1\text{R}_2\text{CN}$ fluorescence and our observed signal-to-noise ratio for OH lines, we can estimate an upper limit to the $\text{R}_1\text{R}_2\text{CN}$ fluorescence quantum yield, provided we assume that the concentrations of $\text{R}_1\text{R}_2\text{CN}$ and OH are equal in the photolysis zone. Ogilvie and Horne⁵ have estimated the integrated oscillator strength for the 280 nm band of H_2CN to be $(4 \pm 2) \times 10^{-4}$. We shall assume that the oscillator strength of the corresponding bands in the radicals produced from acetaldoxime and acetoxime are similar. Since the band is approximately 200 cm^{-1} wide, this implies an average oscillator strength per unit bandwidth of $2 \times 10^{-6}/\text{cm}^{-1}$. We can estimate the oscillator strengths of the OH (1,0) and (2,1) bands from the known radiative lifetimes and fluorescence branching ratios:²²⁻²⁵ $f_{1,0} = 3.1 \times 10^{-4}$ and $f_{2,1} = 4.5 \times 10^{-4}$.

Our most sensitive searches for $\text{R}_1\text{R}_2\text{CN}$ fluorescence were carried out with acetaldoxime and acetoxime and with nitrogen added as a thermalizing buffer. Under these conditions, lines in the OH (2,1) band could be observed with signal-to-noise ratio of greater than 100 in some scans. The oscillator strengths given in the previous paragraph and the previously determined OH ($v=1$) to ($v=0$) population ratio imply that the ratio of an $\text{R}_1\text{R}_2\text{CN}$ signal to that of a low N line in the OH (2,1) band should equal approximately unity if the fluorescence quantum yield ϕ of $\text{R}_1\text{R}_2\text{CN}$ were unity, assuming also that the quantum yield for OH excited fluorescence is also unity. This suggests that ϕ is less than 1% for CH_3CHCN and $(\text{CH}_3)_2\text{CN}$.

The N_2 pressure and delay time of these scans were sufficiently low to prevent vibrational relaxation²⁶ in the OH($X^2\Pi$) state. In fact, the zero-pressure fluorescence quantum yield for OH $v'=2$ is considerably less than unity because of predissociation.^{20,21} The measured radiative lifetime (ca. 450 nsec, which is significantly longer than the lifetime of low N levels in $v'=2$ under collision-free conditions^{20,21}) of the laser excited OH ($v'=2$) indicated that vibrational relaxation²⁷ to the nonpredissociating lower vibrational levels was occurring in the $\text{A}^2\Sigma^+$ state; this would have the effect of substantially raising the

fluorescence quantum yield over that for $v'=2$ in the absence of collisions.^{20,21} Our estimate for the upper bound on ϕ would be even further reduced if the quantum yield of OH fluorescence were less than unity. Hence, we conclude ϕ is certainly less than 1% for CH_3CHCN and $(\text{CH}_3)_2\text{CN}$. The derived upper limit for ϕ is approximately one order of magnitude larger for H_2CN since the OH signals observed for formaldoxime precursor were smaller.

While not observing fluorescence attributable to $\text{R}_1\text{R}_2\text{CN}$, we did observe weak fluorescence from an as yet unidentifiable species other than OH. Figure 3 shows an excitation spectrum of a thermalized sample of photolyzed acetaldoxime in the region of the OH (2,1) band head. In contrast to the nascent spectrum in Figure 2, the high N lines of the OH (1,0) band are very weak compared to those of the (2,1) band because of rotational relaxation in $v''=1$. In addition to the strong OH lines, we can see a weak set of lines with closer spacing which appear to be converging to a band head degraded to the red near 34650 cm^{-1} . In addition to these lines, a second set of more closely spaced lines were observed around 34150 cm^{-1} , as shown in Figure 4. The spacing between lines in the latter region is much smaller than in the former.

We do not believe that these lines are due to $\text{R}_1\text{R}_2\text{CN}$ since (1) they do not match the wavelengths reported for $\text{R}_1\text{R}_2\text{CN}$ from the flash photolysis studies and, more importantly, (2) they appear at identical wavelengths with acetoxime as the precursor. (The lines near 34650 cm^{-1} also appeared, weakly, with formaldoxime.) In an attempt to identify the molecular carrier of these lines, we measured the radiative lifetimes of several of the lines. We estimate the fluorescence decay lifetime to be approximately $1\text{ }\mu\text{sec}$, roughly independent of the line excited and the nitrogen pressure over 0.3-1 Torr. The precision of our measurement is poor because the signal is small compared with the probe laser scattered light. One possible candidate for the molecular species is methoxy, CH_3O , which does have bands in this region²⁸ and whose radiative lifetime is reported to be $1.5\text{ }\mu\text{sec}$.²⁹ However, the rotational structure for a room-temperature sample of methoxy would be expected to be much more dense than the spectra shown in Figures 3 and 4.

At present, we do not have a suitable candidate species to attribute to these bands. We do, however, have several comments to make about the bands. Firstly, there is so much interference from the strong OH lines obscuring the short wavelength, better resolved band that it is difficult even to attempt to fit the rotational structure of the band. The OH lines make it particularly difficult to observe the band origin which would facilitate fitting the spectrum. Secondly, the splitting between the two bands is about 450 cm^{-1} . This mitigates against the possibility that the two bands are subbands of the same vibrational band

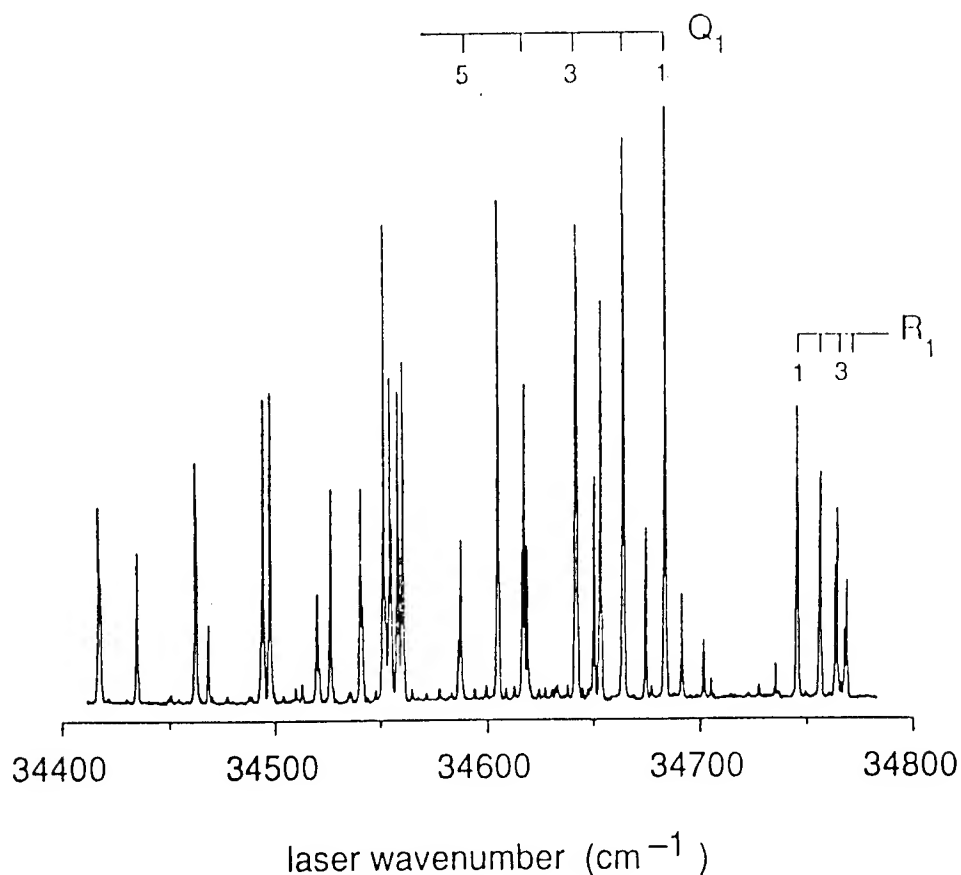


Figure 3. Excitation Spectrum Near the OH A-X (2,1) Band Origin for Photodissociation of Acetaldoxime Under Thermalized Conditions. All the strong lines are due to OH. The Q_1 and R_1 branches of the (2,1) band are marked. The splitting of the R_1 branch lines is due to strong satellite transitions near the main branch lines for these low N values. The pressure and pump-probe delay were 0.66 Torr (nitrogen added) and 25 μ s, respectively.

resulting from spin-orbit splitting because the splitting is so large. If the two bands belong to the same molecule, it would therefore seem that they are different vibrational bands. The form of the short wavelength bands appears to be much simpler than the long wavelength band. The former may consist of only P and R branches, with no apparent

spin splitting, while the latter apparently has more branches. If the bands arise from the same molecular species, it therefore seems likely that the vibrational symmetries of the levels involved in the two bands are different. Alternatively, there could be a K' dependent predissociation in the short wavelength band, as in the $\text{HCO } \tilde{A}^2A'' - \tilde{X}^2A'$ spectrum.^{30,31}

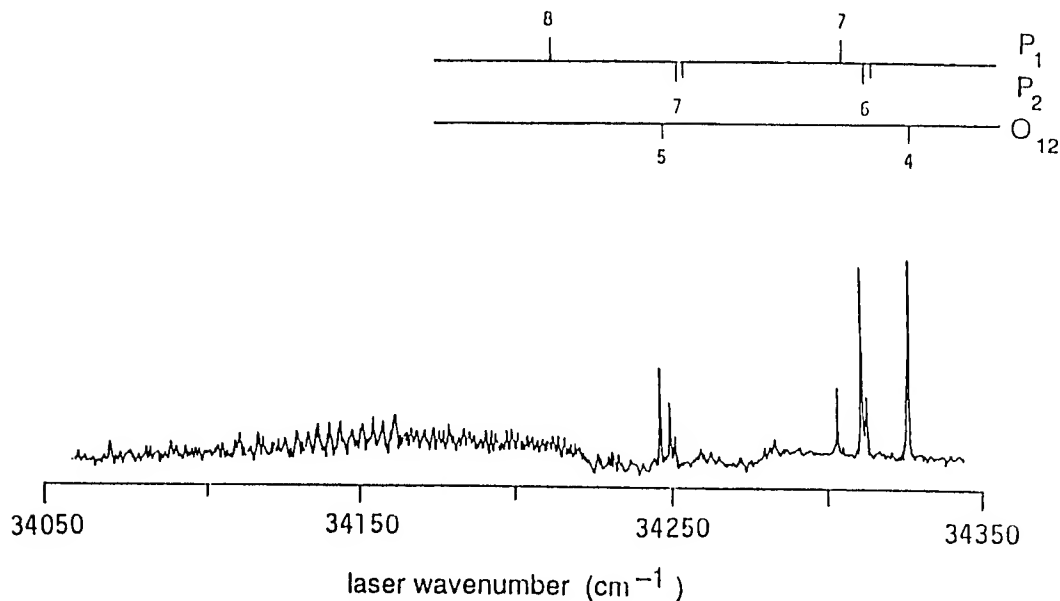


Figure 4. Excitation Spectrum Around 34150 cm^{-1} for Photodissociation of Acetaldoxime Under Thermalized Conditions, Showing a Band of an Unidentified Species. Lines of the OH A-X (2,1) band are marked.

D. Detection of CN: In estimating the upper bound to ϕ , it is possible that our assumption of equal concentrations of the OH and $\text{R}_1\text{R}_2\text{CN}$ photofragments may not be correct. The flash photolysis work⁵⁻⁸ suggests that the latter is not rapidly consumed by chemical reactions since the $\text{R}_1\text{R}_2\text{CN}$ absorptions were seen to persist at much higher reactant pressures than used in the present work for at least $80 \mu\text{sec}$.

However, in our experiments, it is possible, at least in principle, that R_1R_2CN could itself be destroyed in the initial photodissociation since the energy available to the fragments considerably exceeds the R_1-R_2CN bond energy (see Section A) and only a small fraction of this energy appears as OH excitation (see Section B). Thus, both $R_1R_2CN + OH$ and $R_1CN + R_2 + OH$ photolysis products could be formed in a one-photon process. The dissociation wavelength in the present experiment (193 nm) is somewhat shorter than in the flash photolysis studies since the Kr flashlamp output there dropped rapidly below 200 nm.⁷ The effective wavelength range of the photolysis radiation is actually a convolution of the flashlamp output and the oxime absorption coefficient, which is rapidly rising below 200 nm. Hence, it is possible that fragmentation of R_1R_2CN is more likely with photolysis using 193 nm radiation, than with a Kr flashlamp. In fact, the appearance of successive absorption maxima in the continuum spectra of both the oxime and azine precursors^{5,7,8} as one goes toward shorter wavelengths may indicate the onset of some fraction of higher fragmentation photolysis.

In an effort to determine the importance of such destruction of R_1R_2CN , we looked for production of CN by laser fluorescence excitation. Indeed, we observed CN through fluorescence excitation in its $B^2\Sigma^+ - X^2\Sigma^+ (0,0)$ band upon 193 nm photolysis of acetaldoxime. No attempt was made to measure quantitatively the CN concentration relative to that of OH because of the widely different wavelength range, and laser and filter bandwidths. However, we estimate that the CN concentration was significantly less than that of OH since comparable photomultiplier signals were observed; the oscillator strength of the CN $B-X (0,0)$ band equals $(3.11 \pm 0.05) \times 10^{-2}$ ³² and is hence much larger than for the OH ($\Delta v=+1$) sequence. The dependence of the CN fluorescence signal was investigated for several lines as a function of excimer laser pulse energy under thermalizing conditions (1.0 Torr nitrogen added, 4 μ sec pump/probe delay). On a log-log plot, the slope was found to equal 0.78 ± 0.1 . This suggests that CN is formed by a one-photon dissociation process from the oxime. The process $CH_3CHNOH \rightarrow CH_4 + CN + OH$ requires approximately 110 kcal/mole. This dissociation pathway is clearly feasible in a one-photon process at 193 nm.

In addition to the observation of ground state CN by laser fluorescence detection, an emission signal coincident with the ArF excimer laser pulse was seen through the 390 nm bandpass filter with the acetaldoxime precursor. We do not believe this is due to direct photolytic production of $CN(B^2\Sigma^+)$ because the decay time of this signal (ca. 1 μ sec) is much longer than the $CN(B)$ radiative lifetime (ca. 70 nsec⁵¹). The dependence of the emission signal on the excimer pulse energy was found to be approximately linear (slope of 0.99 ± 0.1 on a log-log plot). However, production of $CN(B)$ in a one-photon process at 193 nm is not energetically allowed. Unfortunately, it was not possible

in the present apparatus to take a spectrum of this emission, thus elucidating the species responsible.

IV. CONCLUSION: In this work, we have investigated the 193 nm photodissociation of the simplest oximes, namely formaldoxime, acetaldoxime, and acetoxime. The OH photofragment was detected, and its internal state distribution determined. An unsuccessful search was made to observe laser excited fluorescence from the other photofragment R_1R_2CN , where R_1 and R_2 equals H or CH_3 . Based on the assumption of an equal concentration of OH and R_1R_2CN upon photolysis of the oxime, we were able to estimate a rough upper bound of 1% for the fluorescence quantum yield ϕ of the R_1R_2CN absorption bands in the 280-295 nm region. This implies that electronic states excited in these transitions are predissociative. Through the detection of CN, there is some evidence that the R_1R_2CN is itself destroyed to some extent in the initial event of the 193 nm photodissociation of the oxime. This could have the effect of increasing our estimate on the upper bound to ϕ .

The present study indicates that fluorescence excitation of the electronic bands near 280 nm of methyleneamidogen is not a feasible laser diagnostic for this species. An alternative, more sensitive technique may be resonant-enhanced multiphoton ionization through a Rydberg state, as has been successfully carried out for HCO.^{33,34}

Recently, Marston, et al.,³⁵ have shown that H_2CN can be produced in high yield from the $N + CH_3$ reaction. Using the discharge flow/mass spectrometry technique, this group has measured the elementary reaction rate constant for the $N + H_2CN$ reaction. These results imply that mass spectrometric detection of H_2CN may be a viable alternative diagnostic for this species. In addition, it would be interesting to compare the production of H_2CN by this chemical method with our photolytic approach.

ACKNOWLEDGMENT: We appreciate the assistance of M.A. Schroeder in the synthesis of several of the precursors. We also acknowledge M.A. Schroeder, R.A. Fifer, and K.L. McNesby for their aid in identification of these precursors. We thank B.J. Gaffney for the loan of a distillation apparatus used in the syntheses. P.J.D. acknowledges the support of the U.S. Army Research Office under contract number DAAL-03-88-K-0031 and the National Science Foundation under grant number CHE-8700970. A.W.M. acknowledges partial support from the Air Force Office of Scientific Research, Directorate of Aerospace Sciences, under contract number 89-0017.

REFERENCES

1. D.R. Crosley, Ed., "Laser Probes for Combustion Chemistry", Amer. Chem. Soc. Symposium Series; Amer. Chem. Soc., Washington, DC, Vol. 134, 1980.
2. J.H. Bechtel, C.J. Dasch, R. Teets, In Laser Applications, J.F. Ready and R.K. Erf, Eds.; Academic Press, New York, Vol. 5, p 129, 1983.
3. (a) D.R. Crosley, "Collisional Effects on Laser-Induced Fluorescence Flame Measurements", Opt. Eng., Vol. 20, p. 511, 1981;
(b) D.R. Crosley and G.P. Smith, "Laser-Induced Fluorescence Spectroscopy for Combustion Diagnostics", Opt. Eng., Vol. 22, p. 545, 1983.
4. E.L. Cochran, F.J. Adrian, and V.A. Bowers, "ESR Detection of the Cyanogen and Methylene Imino Free Radicals", J. Chem. Phys., Vol. 36, p. 1938, 1962.
5. J.F. Ogilvie and D.G. Horne, "Electronic Absorption Spectra of Methanal Azine and the Methyleniminyl Free Radical", J. Chem. Phys., Vol. 48, p. 2248, 1968.
6. J.F. Ogilvie, "Assignments of Electronic Transitions in Methanal Azine and Related Molecules", Can. J. Spectrosc., Vol. 19, p. 89, 1974.
7. D.G. Horne and R.G.W. Norrish, "The Photolysis of Acyclic Azines and the Electronic Spectra of $R_1R_2CN_7$ Radicals", Proc. R. Soc. London A, Vol. 315, p. 301, 1970.
8. D.G. Horne and R.G.W. Norrish, "The Flash Photolysis of Oximes", Proc. R. Soc. London A, Vol. 315, p. 287, 1970.
9. R.A. Fifer, "Chemistry of Nitrate Ester and Nitramine Propellants", In Fundamentals of Solid-Propellant Combustion; K.K. Kuo and M. Summerfield, Eds., Prog. in Aeronautics and Astronautics, Vol. 90, p. 177, 1984.
10. C.F. Melius and J.S. Binkley, "Thermochemistry of the Decomposition of Nitramines in the Gas Phase", Proceedings of the 21st International Symposium on Combustion, Combustion Institute, Pittsburgh, PA, p. 1953, 1986.
11. C.U. Morgan and R.A. Beyer, "Electron-Spin-Resonance Studies of HMX Pyrolysis Products", Combust. Flame, Vol. 36, p.99, 1979.

12. X. Zhao, E. Hints, and Y.T. Lee, "Infrared Multiphoton Decomposition of RDX in a Molecular Beam", J. Chem. Phys., Vol. 88, p. 801, 1988.
13. P.J. Dagdigian, "State Resolved Cross Sections for Rotationally Inelastic Collisions of $\text{NH}_2(\tilde{\text{A}}_1)$ with Helium", J. Chem. Phys., Vol. 90, p. 2617, 1989.
14. R. Scholl, "Das Einfachste Normale Oxim $\text{H}_2\text{C:NOH}$ und Sein Polymeres", Berichte, Vol. 24, p. 573, 1891.
15. S.W. Benson and H.E. O'Neal, "Kinetic Data on Gas Phase Unimolecular Reactions", Natl. Stand. Ref. Data Ser., U.S., Natl. Bur. Stand., p. 21, 1970.
16. R.A. Bair and T.H. Dunning, Jr. "Theoretical Studies of the Reactions of HCN with Atomic Hydrogen", J. Chem. Phys., Vol. 82, p. 2280, 1985.
17. M.W. Chase, Jr., C.A. Davies, J.R. Downey, Jr.; D.J. Frurip, R.A. McDonald, and A.N. Syverud, "JANAF Thermochemical Tables, 3rd Edition", J. Phys. Chem. Ref. Data, Vol. 14, Supplement 1, 1985.
18. M. McKee and C.F. Melius, Private communication.
19. C.H. Greene and R.N. Zare, "Determination of Product Population and Alignment Using Laser-Induced Fluorescence", J. Chem. Phys., Vol. 78, p. 6741, 1983.
20. K.R. German, "Radiative and Predissociative Lifetimes of the $v'=0,1$, and 2 Levels of the $\text{A}^2\Sigma^+$ State of OH and OD", J. Chem. Phys., Vol. 63, p. 5252, 1975.
21. J. Brzozowski, P. Erman, and M. Lyyra, "Precision Estimates of the Predissociation Rates of the OH $\text{A}^2\Sigma$ State ($v' \leq 2$)", Phys. Scripta, Vol. 17, p. 507, 1978.
22. W.L. Dimpfl and J.L. Kinsey, "Radiative Lifetimes of OH ($\text{A}^2\Sigma$) and Einstein Coefficients for the A-X System of OH and OD", J. Quant. Spectrosc. Radiat. Transfer, Vol. 21, p. 233, 1979.
23. I.S. McDermid and J.B. Laudenslager, "Radiative Lifetimes and Quenching Rate Coefficients for Directly Excited Rotational Levels of OH ($\text{A}^2\Sigma^+ v'=0$)", J. Chem. Phys., Vol. 76, p. 1824, 1982.

24. D.R. Crosley and R.K. Lengel, "Relative Transition Probabilities in the A-X System of OD", J. Quant. Spectrosc. Radiat. Transfer, Vol. 17, p. 59, 1977.
25. R.A. Copeland, J.B. Jeffries, and D.R. Crosley, "Transition Probabilities in OH $A^2\Sigma^+ - X^2\Pi$: Bands with $v'=0$ and 1, $v''=0$ to 4", Chem. Phys. Lett., Vol. 138, p. 425, 1987.
26. K.J. Rensberger, J.B. Jeffries, and D.R. Crosley, "Vibrational Relaxation in OH ($X^2\Pi$, $v=2$)", J. Chem. Phys., Vol. 90, p. 2174, 1989.
27. R.A. Copeland, M.L. Wise, and D.R. Crosley, "Vibrational Energy Transfer and Quenching of OH ($A^2\Sigma^+$, $v'=1$)", J. Phys. Chem., Vol. 92, p. 5710, 1988.
28. S.C. Foster, P. Misra, T.Y.D. Lin, C.P. Damo, C.C. Carter, and T.A. Miller, "Free Jet Cooled Laser-Induced Fluorescence Spectrum of Methoxy. 1. Vibronic Analysis of the \tilde{A} and \tilde{X} States", J. Phys. Chem., Vol. 92, p. 5914, 1988.
29. G. Inoue, J. Akimoto, and M. Okuda, "Laser-Induced Fluorescence Spectra of CH_3O ", Chem. Phys. Lett., Vol. 63, p. 213, 1979.
30. J.W.C. Johns, S.H. Priddle, and D.A. Ramsay, "Electronic Absorption Spectrum of HCO and DCO Radicals", Discuss. Faraday Soc., Vol. 35, p. 90, 1963.
31. R. Vasudev and R.N. Zare, "Laser Optogalvanic Study of HCO \tilde{A} State Predissociation", J. Chem. Phys., Vol. 76, p. 5267, 1982.
32. N. Duric, P. Erman, and M. Larsson, "The Influence of Collisional Transfers and Perturbations on Measured A and B State Lifetimes in CN", Phys. Scripta, Vol. 18, p. 39, 1978.
33. P.J.H. Tjossem, P.M. Goodwin, and T.A. Cool, "Two-Photon Resonance REMPI Detection of the Formyl Radical", J. Chem. Phys., Vol. 84, p. 5334, 1986.
34. J.S. Bernstein, X.M. Song, and T.A. Cool, "Detection of the Formyl Radical in a Methane/Oxygen Flame by Resonance Ionization", Chem. Phys. Lett., Vol. 145, p. 188, 1988.
35. G. Marston, F.L. Nesbitt, and L.J. Stief, "Branching Ratios in of N Atoms with N + CH_3 Reaction: Formation of the Methyleneamidogen (H_2CN) Radical", J. Chem. Phys., Vol. 91, p. 3483, 1989.

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

Expression of Hantaan Viral Antigens for Vaccine Development

Connie S. Schmaljohn, Dr., Yong-Kyu Chu, Dr., and *Joel M. Dalrymple, Dr.
United States Army Medical Research Institute of Infectious Diseases
Ft. Detrick, Frederick, MD. 21701-5011

Korean hemorrhagic fever (KHF) first gained prominence in 1951 when an explosive epidemic occurred among United Nations forces stationed in Korea. A mortality rate of 10-15% was reported for approximately 3,000 soldiers hospitalized with KHF during the Korean conflict. Since that time, KHF has remained endemic in Korea, and continues to pose a serious health threat to troops in that region. Moreover, KHF is only one of many similar diseases that have been clinically diagnosed throughout Europe and Asia. KHF-like diseases have been reported in China, Japan, USSR, Czechoslovakia, Romania, Hungary, Bulgaria, Yugoslavia, Finland, Sweden, Norway, Denmark, Belgium, France, Greece, and West Germany. More than 150 synonyms have been used to describe these diseases; however, the World Health Organization has established the collective title of "hemorrhagic fever with renal syndrome" (HFRS).

Although HFRS has been recognized clinically for many years, the first glimpse at its etiological origin did not occur until 1978, when Hantaan virus was identified in the lungs of Korean striped field mice¹. Biochemical characterization of Hantaan virus by our laboratory provided the first indisputable evidence that Hantaan virus is molecularly similar to members of the *Bunyaviridae* family of animal viruses. We found that, like other viruses in the family, Hantaan virus has a three-segmented, single-stranded RNA genome of anti-message sense, enclosed in three ribonuclease-sensitive nucleocapsids surrounded by a lipid envelope containing two virus-specified glycoproteins^{2,3,4,5}. No serological relationship, however, could be demonstrated between Hantaan virus and any other member of the *Bunyaviridae*. Only after molecular characterization of several other virus isolates, which were serologically related to Hantaan virus yet antigenically unique, was it established that Hantaan and related viruses comprise a new, and separate genus of *Bunyaviridae*⁴. Our proposal to establish the *Hantavirus*

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

genus was formally accepted by the International Committee on the Taxonomy of Viruses in 1987.

It is now clear that hantaviruses exist in indigenous rodent populations throughout the world. Isolates, which are nearly indistinguishable from one another by serological means, have been obtained both from endemic disease regions and also in geographic areas where typical HFRS has not been reported, including the United States and Canada. The need for a safe and effective vaccine to prevent HFRS remains of tantamount importance both for military and civilian personnel who must travel or live in areas where hantaviruses are circulating. Traditionally, vaccines to viruses are produced either by attenuation or inactivation of virulent viruses. A live-attenuated Hantaan virus vaccine has not been attempted, and may be unwise, because of the propensity of these viruses to cause persistent infections in animals. Efforts to develop killed vaccines have been hindered by the slow and low-titered replication of the viruses in cell culture and the inability to propagate them in most certified cell lines. Also of concern are the laboratory hazards associated with producing large quantities of these aerosol-transmissible viruses. These factors can all be circumvented by the use of a recombinant DNA vaccine for HFRS. To this end, we have prepared two potential vaccines for HFRS by inserting Hantaan viral genes into two eucaryotic viruses, one of which is infectious and the other non-infectious for humans.

RESULTS

Cloning and sequence analysis of Hantaan viral genes. The first step toward development of a recombinant DNA vaccine for HFRS was to produce cDNA clones representing the Hantaan viral RNA and to determine their nucleotide sequences. Hantaan virus has three genome segments: small (S), medium (M) and large (L). If analogous to other viruses in the *Bunyaviridae* family, the S segment would encode the viral nucleocapsid protein (N), the M segment the two envelope glycoproteins (G1 and G2), and the large segment the virion-associated transcriptase. We were most interested in the S and the M genome segments because the structural proteins they encode (particularly the outer envelope proteins) would be expected to be the targets of any protective immune response to Hantaan virus. Cloning and sequence analysis revealed that the Hantaan viral S genome segment consisted of 1696 nucleotides, and contained only one, long, open-reading frame (ORF), located in the virus-complementary sense RNA, which could encode a protein of 48

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

kD (Fig.1). To prove that this ORF encoded N, we used SP6 polymerase to generate RNA from the cDNA *in vitro*. Translation of that RNA in a cell-free system resulted in a protein that could be immune-precipitated with antibodies to authentic Hantaan viral nucleocapsid protein, thus demonstrating that the S segment of Hantaan encodes N⁶.

The Hantaan M genome segment was found to consist of 3616 bases and, again, a single ORF was detected in the virus-complementary sense RNA. This reading frame encoded a potential polypeptide of approximately 126,000 daltons, which is of sufficient size to include both G1 and G2 (Fig. 1). To prove that the M segment encoded the two envelope glycoproteins, we purified authentic Hantaan virus G1 and G2 and sequenced their amino-terminal amino acids. We then compared the sequences to those deduced from the nucleotide sequences of the M genome segment. From this, we were able to determine that Hantaan M encodes G1 and G2 in a single, continuous ORF and the gene order is 5'-G1-G2-3' with respect to the message-sense RNA (Fig. 1)⁷.

Expression of Hantaan viral genes. Two eucaryotic viruses were chosen as expression vectors for Hantaan viral genes: the insect baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) and vaccinia virus. Because baculoviruses cannot infect mammals, the AcNPV system offers the advantages of a traditional killed vaccine without the potential risk of incomplete inactivation of infectious virus. Conversely, because vaccinia virus does infect humans, it offers the advantages of a replicating vaccine: *i.e.*, a small amount of vaccine can produce sufficient antigen to generate a protective immune response. Thus the vaccinia virus system offers the advantages of an attenuated vaccine, without the risk of reversion to the virulent strain of virus, which could cause HFRS. Both the baculovirus and vaccinia virus expression systems offer distinct advantages over authentic Hantaan virus for vaccine development. For example, they can both be used under biosafety level 2 containment conditions, whereas Hantaan virus requires biosafety level 3. Also, both replicate rapidly and to high titer, whereas Hantaan virus replicates slowly and to low titer and causes persistent infections of the few cell lines in which it will grow.

To prepare recombinant viruses expressing Hantaan viral N, G1 and G2 we first inserted cDNA representing the entire S or M segments or restriction fragments representing portions of the M segment encoding only G1 or only

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

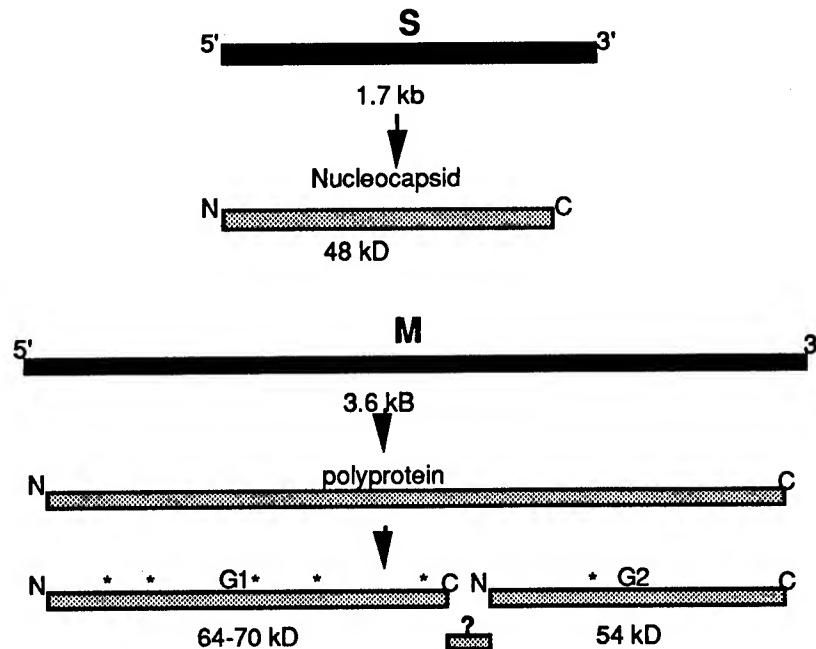


Fig. 1. The small (S) and the medium (M) genome segments of Hantaan virus were molecularly cloned and nucleotide sequences were determined by dideoxy chain termination reactions. The small genome segment consists of 1696 nucleotides (1.7 kB) and encodes the 48,000 dalton (48 kD) viral nucleocapsid protein in the virus-complementary sense RNA. The M genome segment consists of 3616 nucleotides (3.6 kB) and encodes G1 and G2 in a single continuous open-reading frame, also in the virus-complementary sense RNA. Five potential asparagine-linked glycosylation sites (*) were detected in the deduced amino acid sequence of G1 and one in G2.

G2, into plasmid transfer vectors, which also contained sequences homologous to portions of the thymidine kinase gene of vaccinia virus or the polyhedrin gene of AcNPV. Homologous recombination between these plasmid DNAs and vaccinia or baculovirus DNA was performed in cell cultures according to previously published techniques^{8,9,10} and resulted in the Hantaan viral genes being incorporated into the thymidine kinase gene of vaccinia virus (under control of the vaccinia 7.5 K promoter) or into the polyhedrin gene of AcNPV (under control of the polyhedrin promoter). Selection of baculovirus

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

recombinants was accomplished by limiting dilution and DNA hybridization. The plasmid used to transfer Hantaan genes into vaccinia virus also contained the Lac Z gene of the bacteria *Escherichae coli*. Recombinants expressing the Hantaan genes, therefore, also expressed the Lac Z gene. Because the Lac Z gene product, beta-galactosidase, forms a blue color in the presence of the substrate X-gal, it was possible to directly select blue virus plaques (recombinant vaccinia viruses) from white plaques (non-recombinants) by assay on cell-culture monolayers.

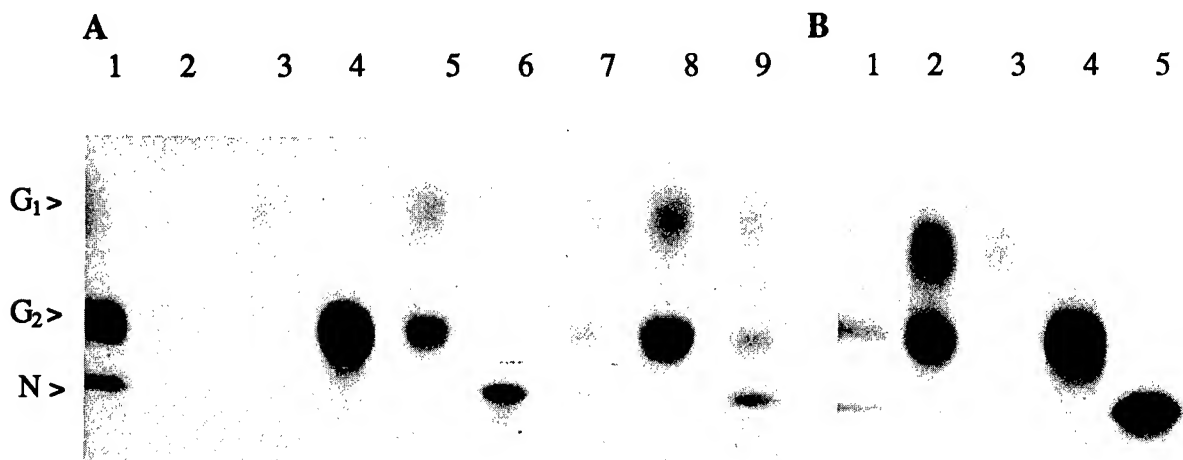


FIG. 2. Polyacrylamide gel electrophoresis of radiolabeled proteins immune-precipitated from infected cell lysates. A. Authentic Hantaan viral proteins were precipitated from Vero E6 cell lysates with anti-Hantaan hyperimmune mouse ascitic fluid (lane 1) or normal mouse ascitic fluid (lane 2). Proteins generated by recombinant vaccinia virus expression of the following Hantaan genes were precipitated from Vero E6 cell lysates with anti-Hantaan hyperimmune mouse ascitic fluid: lane 3, G1; lane 4, G2; lane 5, coinfection with recombinants expressing only G1 and only G2; lane 6, S; lane 7, coinfection with recombinants expressing only G1, only G2 and S; lane 8, M; lane 9, coinfection with recombinants expressing M or S. B. Authentic Hantaan proteins (lane 1) were precipitated from infected Vero E6 cell lysates with anti-Hantaan hyperimmune mouse ascitic fluid. Recombinant baculoviruses expressing the following genes were precipitated from infected *Spodoptera frugiperda* cell lysates with anti-Hantaan hyperimmune mouse ascitic fluid: lane 2, M; lane 3, G1; lane 4, G2; lane 5, S. The positions of Hantaan envelope proteins (G1, G2) and nucleocapsid protein (N) are indicated with arrows.

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

To examine the expression products, cell cultures infected with recombinant baculoviruses or vaccinia viruses were radiolabeled with ^{35}S -cysteine and Hantaan-specific antisera were used to immune-precipitate proteins from cell lysates. Proteins were then subjected to polyacrylamide gel electrophoresis and visualized by autoradiography. Expressed proteins from either recombinant baculoviruses or vaccinia viruses were indistinguishable from authentic Hantaan virus N, G1 and G2 by this method (Fig. 2).

Antigenicity of the expressed proteins. To determine whether the expressed proteins retained antigenic properties similar to those of authentic Hantaan viral proteins, cell cultures infected with recombinant viruses were screened by an indirect immunofluorescent antibody test with monoclonal antibodies (MAbs) to the G1 or G2 proteins of Hantaan virus¹¹. Six MAbs to G1, which recognize two distinct antigenic sites, and 14 MAbs to G2, which recognize seven distinct sites, were tested against SF-9 cells infected with the baculovirus recombinants, or against Vero E6 cells infected with the vaccinia virus recombinants. All of the G1-specific MAbs reacted with proteins from recombinants expressing G1, or those expressing both G1 and G2, but not with those expressing only G2, N, or with uninfected cell controls. Similarly, the anti-G2 MAbs reacted only with cells infected with recombinants expressing G2, or both G1 and G2. These data indicated that at least those antigenic sites defined by these MAbs were conserved on the expressed proteins.

Immunogenicity of expressed proteins. No animal model has yet been discovered that mimics the symptoms of HFRS. In order to determine if the expressed proteins could induce a protective immune response, we developed an infectivity model in Syrian hamsters. As listed in Table 1, we found that, with a Hantaan virus dose of 1000 plaque forming units (PFU), administered intramuscularly, all hamsters were infected and displayed antigen in their lungs. Lower doses of virus did not infect all animals consistently, and higher doses resulted in less detectable antigen in the animals' lungs, possibly reflecting antigen clearance by 28 days post infection. Therefore, for all protection experiments, animals were challenged with 1000 PFU of Hantaan virus.

The presence of antibodies capable of neutralizing a virus in cell culture is generally thought to correlate with a protective immune response. To examine the participation of specific viral epitopes in the protective immune response, and to determine if neutralizing antibody to Hantaan virus correlated with

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

TABLE 1. Hamster challenge model

Virus Dose	Infected/ inoculated ^a	Antigen in lungs ^b	ELISA titer ^c
10 PFU	3/5	+++;++	320
100 PFU	4/5	+,+++	360
1000 PFU	5/5	++,+++	4830
10000 PFU	5/5	+,+++	5120
100000 PFU	5/5	0, +	>12800

^a as determined by detectable antibody response.^bLowest (+) and highest (+++) amount of antigen detected in lungs.^cAverage titer of all infected animals in the group (28 days post-challenge).

protection, we passively transferred anti-Hantaan MAbs¹¹ recognizing seven different sites on G2 and two on G1 to hamsters and subsequently challenged them with Hantaan virus. Most of the animals receiving neutralizing G2 MAbs did not develop antibodies detectable by immunofluorescent antibody staining after challenge, and only one animal displayed antigen in its lungs (Table 2). These results suggest that the MAbs protected the hamsters from infection. Although the three G1-specific neutralizing MAb recognize the same or closely related antigenic sites (as revealed by competitive inhibition studies)¹¹, one of these antibodies appeared to protect animals completely from infection, while the other two appeared to offer only limited protection. None of the non-neutralizing antibodies or normal mouse ascitic fluid appeared to prevent animals from developing high-titer antibody responses after challenge, and most animals displayed antigen in their lungs, which together indicate that animals were infected. These data suggest that a neutralizing antibody response alone is sufficient to protect hamsters from infection with Hantaan virus, and that neutralizing antibodies to either G1 or G2 can passively confer protection.

To evaluate the ability of the recombinant-expressed proteins to elicit a protective response in hamsters, animals were given two immunizations with each recombinant virus preparation. Serum antibody titers were measured after each immunization and again after challenge. Animals were killed 28 days after challenge and lung sections examined for the presence of Hantaan

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

viral antigen. The hamsters receiving the baculovirus or vaccinia virus recombinants expressing the complete M segment developed neutralizing antibody titers before challenge and did not display antigen in their lungs after challenge (Table 3). Most animals immunized with the G1, G2 or M segment recombinants did not display antibodies detectable by an indirect immunofluorescent antibody test (IFAT) prior to challenge; therefore, the presence of IFAT-detectable antibodies after challenge suggested that the animals became infected with the Hantaan virus challenge.

TABLE 2. Passive Protection of Hamsters with Hantaan Monoclonal Antibodies

MAb ^a	Serum antibody ^{b,c}	Antigen in lungs ^c
<u>Neut. MAb</u>		
G2a (HCO2)	1 / 4	0 / 4
G2a (16E6)	1 / 4	1 / 4
G2c (11E10)	1 / 4	0 / 4
G1b (3D5)	0 / 4	0 / 4
G1b (2D5)	4 / 4	1 / 4
G1b (16D2)	4 / 4	2 / 4
<u>Non-neut Mab</u>		
G2b (EBO6)	4 / 4	4 / 4
G2d (3D7)	4 / 4	3 / 4
G2e (20D3)	4 / 4	3 / 4
G2f (23G10)	3 / 3	3 / 3
G2f (GDO5)	4 / 4	3 / 4
G2f (8E10)	4 / 4	3 / 4
G2g (JDO4)	4 / 4	4 / 4
G1a (6D4)	3 / 3	3 / 3
G1a (8B6)	4 / 4	2 / 4

^aDistinct epitopes on G1 and G2 (G1 a-b and G2 a-f.) were previously defined by competitive inhibition.¹¹

^btiters \geq 1:64 at 21 days post-challenge

^cnumber +/total

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

No post-challenge, Hantaan-specific antibodies were observed in the sera of animals immunized with the complete M segment recombinant preparations, suggesting that they were protected from infection. Similar results were obtained in hamsters immunized with control preparations of authentic Hantaan virus, which were inactivated by irradiation (Table 3). In contrast, animals immunized with preparations of cells infected with baculovirus recombinants expressing only G1 or G2 coding regions did not

TABLE 3. Challenge of Immunized Hamsters With Virulent Hantaan Virus

Immunogen	Pre-challenge				Post-Challenge	
	PRNT ^b		IFAT ^c		Antigen in Lungs	IFAT ^c Antibody
	Antibody		Antibody			
	1°	2°	1°	2°		
Baculo-M	4/5	5/5	2/5	2/5	0/5	0/5
Baculo-G1	0/5	0/5	0/5	0/5	0/5	2/5
Baculo-G2	0/5	0/5	0/5	0/5	0/5	4/5
Baculo-S	0/5	0/5	5/5	5/5	0/5	5/5
Uninfected	0/5	0/5	0/5	0/5	4/5	4/5
Cell Control						
Vac-WR-M	5/5	5/5	0/5	0/5	0/5	0/5
Vac-Con-M	2/5	5/5	0/5	0/5	0/4	0/4
Vac-Con-G1	0/5	0/5	0/5	0/5	1/5	5/5
Vac-Con-G2	0/5	0/5	0/5	0/5	2/5	2/5
Vac-Con-S	0/5	0/5	0/5	0/5	3/5	5/5
HTN virus ^d						
Cell supernate	0/5	5/5	4/5	4/5	0/5	0/5
Purified virus	0/10	7/10	7/10	7/10	0/9	0/9
Infected cells	0/5	2/5	0/5	0/5	1/5	1/5
Uninoculated						
Control ^d	0/5	0/5	0/5	0/5	5/5	5/5

^aValues listed are #positive/#inoculated.

^bPlaque-reduction neutralization test (PRNT).

^cImmunofluorescent antibody test (IFAT).

^dInfectivity inactivated by irradiation with ⁶⁰Co.

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

develop neutralizing antibodies before challenge, and some of the animals became infected after challenge, as indicated by IFAT-detectable serum antibodies. However, none of the animals displayed detectable antigen in their lungs. Similar results were obtained with the vaccinia recombinants producing only G1 or only G2; *i.e.*, hamsters did not develop neutralizing antibody responses and some of the animals became infected and displayed Hantaan viral antigens in their lungs.

Although antibodies to the N protein of Hantaan virus are not presumed to be involved in neutralization, animals immunized with preparations of cells infected with the baculovirus recombinant expressing the S segment did not have antigen in their lungs after challenge. Similar apparent protection was not observed with vaccinia virus recombinants expressing S. Although no attempt was made to quantitate the exact amount of antigen each animal received (and in fact it would be impossible to do so with the replicating recombinant vaccinia viruses), it is likely that the large amount of expressed N produced by the recombinant baculovirus, as compared to that produced by the recombinant vaccinia virus, provided a greater immunizing antigen dose. These data suggest that a non-neutralizing immune response, perhaps a cell-mediated response, can also protect animals from infection with Hantaan virus.

In order to confirm our initial observations that the absence of antigen in the lungs of hamsters, and our inability to detect antibodies by IFAT after challenge indicated protection, we used sera from immunized hamsters to immune-precipitate authentic Hantaan viral proteins. Animals which were protected from infection should have displayed antibodies only to the protein with which they were immunized both before and after challenge, while those not protected should have developed antibodies to all three Hantaan structural proteins (N, G1 and G2) after challenge with Hantaan virus. As before, both vaccinia and baculovirus recombinants expressing the complete M segment were able to induce a protective response in the hamsters, as indicated by antibody responses only to G1 and G2, but not N. Likewise, some of the animals immunized with recombinants expressing individual G1 or G2 were protected and displayed antibodies only to the protein with which they were immunized both before and after challenge, but others developed antibodies to all three proteins after challenge, indicating infection with the challenge virus. As we observed in our previous experiments, two immunizations with preparations containing baculovirus-expressed N protected all animals in that no Hantaan antigen could be detected in lungs, and only N-specific antibodies were observed both before and after challenge.

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

These data support the conclusion that the recombinant proteins are able to induce an immune response in hamsters, which protects them from subsequent infection with Hantaan virus.

DISCUSSION

Advances in molecular biology have afforded opportunities to genetically engineer efficacious vaccines to dangerous etiologic agents with minimal risk to both researchers and vaccine recipients. The genes of deadly viruses, which encode proteins known to be involved in protective immune responses, can be molecularly cloned and inserted into more benign viral hosts. These recombinant viruses will then express not only their own proteins, but also those specified by the foreign gene. We have prepared recombinant viruses containing genes from Hantaan virus, the etiologic agent of Korean hemorrhagic fever, and have evaluated their ability to elicit a protective immune response in animals.

The expressed proteins were found to closely resemble authentic Hantaan proteins and were antigenically identical to those proteins within the limits of our panel of monoclonal antibodies. Moreover, the expressed proteins were able to induce antibodies in test animals, which could both recognize authentic Hantaan viral proteins in serological assays and could neutralize infectious virus.

The absence of HFRS-like disease in animals made it difficult to evaluate the ability of our potential vaccines to moderate disease; however, we were able to use a much more stringent test of experimental vaccine efficacy of the recombinant-expressed proteins: that is, protection from Hantaan virus infection. As demonstrated by passive protection studies, a humoral response to either of the envelope glycoproteins was sufficient to protect against challenge, suggesting that immunization with expressed G1 or G2 proteins might similarly induce a protective immune response. However, we found that immunization with recombinants expressing the entire M segment (*i.e.*, both G1 and G2) was much more effective at eliciting anti-Hantaan antibodies and protecting animals from infection with Hantaan virus than was immunization with those expressing either G1 or G2 genes separately. Whether this is due to an additive effect of antibodies to both proteins or to an interaction of the two proteins which resulted in the formation or stabilization of conformationally-dependent antigenic sites, was not resolved.

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

Previous reports indicated that a cell-mediated response to Hantaan virus may also be involved in protection, but the Hantaan viral component(s) eliciting the response were not identified¹². In our studies, the ability of the expressed nucleocapsid protein to protect animals from Hantaan virus challenge suggests that this protein can induce a non-neutralizing, protective immune response that is, perhaps, cell-mediated. Additional information is required to define the important aspects of immunity to Hantaan virus; however, our data suggest that whatever the mechanism(s), our recombinant-expressed proteins are clearly capable of inducing a response that can protect animals from infection with Hantaan virus and that a humoral response alone is sufficient to provide protection.

The primary goal of studies such as ours is the development of a vaccine to prevent HFRS. Although both the baculovirus and vaccinia virus expression systems yielded products effective in inducing immune responses in animals, each has advantages and disadvantages for human use. We currently are pursuing vaccine development efforts with both expression systems. We have engineered the vaccinia recombinants to eliminate the superfluous expression of the Lac Z gene, which was initially included to aid in the selection of the recombinants. Based on our results, which suggest that expressed N can also protect animals from infection, we have also engineered a potential vaccinia vaccine that expresses both the M and the S segments simultaneously. These vaccines are now being produced in certified cell lines, after which, they will be tested for safety and efficacy in human volunteers.

Because the baculovirus recombinants cannot be propagated in certified mammalian cell lines, a potential problem for their use in humans is the need to recover and purify the expressed proteins from infected insect cell cultures. We have performed preliminary experiments to determine the best methods for purification and further testing awaits preparation of sufficient quantities of the purified products.

CONCLUSION

The data presented here demonstrate the feasibility of using expressed Hantaan proteins to immunize animals to Hantaan virus. These studies should provide a basis for the future exploitation of the recombinant-expressed Hantaan viral proteins as potential human vaccines.

UNCLASSIFIED

UNCLASSIFIED

SCHMALJOHN, CHU, DALRYMPLE

REFERENCES

1. Lee, H.W., P.W. Lee, and K.M. Johnson. (1978) J. Infect. Dis. 137: 298-308.
2. Schmaljohn, C.S., S.E. Hasty, S.A. Harrison, and J.M. Dalrymple. (1983). J. Infect. Dis. 148: 1005-1012.
3. Schmaljohn, C.S., and J.M. Dalrymple. (1983) Virology 131: 482-491.
4. Schmaljohn, C.S., S.E. Hasty, J.M. Dalrymple, J.W. LeDuc, H.W. Lee, C.-H. von Bonsdorff, M. Brummer-Korvenkontio, A. Vaheri, T.F. Tsai, H.L. Regnery, D. Goldgaber, and P.W. Lee. (1985) Science 227: 1041-1044.
5. Schmaljohn, C.S., S.E. Hasty, L. Rasmussen, and J.M. Dalrymple. (1986) J. Gen. Virol. 67: 707-717.
6. Schmaljohn, C.S., G.B. Jennings, J. Hay, and J.M. Dalrymple. (1986) Virology 155: 633-643.
7. Schmaljohn, C.S., A.L. Schmaljohn, and J.M. Dalrymple. (1987) Virology 157: 31-39.
8. Schmaljohn, C.S., K. Sugiyama, A.L. Schmaljohn, and D.H.L. Bishop. (1988) J. Gen. Virol. 69: 777-786.
9. Schmaljohn, C.S., J. Arikawa, J.M. Dalrymple, and A.L. Schmaljohn. (1989) *In* Genetics and pathogenicity of negative strand viruses (D. Kolakofsky and B. Mahy, Eds.), pp 58-66. Elsevier Biomedical Press.
10. Summers, M.D., and G.E. Smith. (1986) Texas Agricultural Experiment Station Bulletin No. 1555.
11. Arikawa, J., A.L. Schmaljohn, J.M. Dalrymple, and C.S. Schmaljohn. (1989) J. Gen. Virol. 70: 615-624.
12. Asada, H., M. Tamura, K. Kondo, Y. Dohi, and K. Yamanishi. (1988) J. Gen. Virol. 69: 2179-2188.

UNCLASSIFIED

KNOWLEDGE WORKER SYSTEM
An Investment for the Army's Future (U)

*Wayne Schmidt, Mr.
Beverly Coskunoglu, Ms.
US Army, Corps of Engineers
Construction Engineering Research Laboratory
Champaign, Illinois 61824-4005

You are an Army Officer.

You just arrived at your new assignment.

Day 1

Your desk is vacant and so is your understanding of the new assignment!

Day 2

The situation is becoming clearer; the phone is ringing and people now know "you are there."

Day 3

The situation is now crystal clear!

You are behind.

The phone rings constantly.

You have 3 overlapping meetings.

And the General wants that report -- due last month -- NOW!

This situation is common throughout the Army. Personnel move to new positions where their assigned tasks are novel and complex. Individual tasks may be automated but WHAT to do and WHEN is largely a function of institutional memory. The Knowledge Worker System captures this institutional memory and provides a performance support environment for these personnel.

The Problem

The Army, like other complex and highly distributed organizations, depends upon a sound infrastructure. "Infrastructure" refers to those embedded services that allow the organization to perform its central mission and function. The central mission of the Army is war fighting but the vast majority of its resources fund support activities. Without its logistics systems, personnel, training and transportation, the Army could not field a fighting force.

The primary effector in the infrastructure is the action officer. Action officers define requirements, allocate resources, review programs, analyze data, provide guidance, gather and disseminate information. They do not perform the Army's primary function but rather provide essential support services. Action officers deal not in hardware or product but rather in *information*. They keep the information flowing and thus, the infrastructure effective. Action officers are therefore "knowledge workers," i.e., professionals who add value to a process or physical product by the results of mental exertions in concert with the leveraging power of automated support systems.

A prime problem facing knowledge workers is keeping track of the multiple tasks. A task, in itself, would be manageable if the knowledge worker could remember what he/she did last year and why. Traditionally, organizations deal with this problem via printed documents: policies, procedures, standard operation processes, regulations and guidelines. These documents attempt to specify WHO is to do WHAT, WHEN and HOW. In large organizations, like the Army, the regulations attempt to insure consistent performance across widely dispersed organizational units. Regulations purport to print institutional memory; however, printed documents are, by their nature, limited in their distribution, currency and specificity.

The Challenges

The current infrastructure presents the knowledge worker with a variety of challenges:

- Increasing number and complexity of tasks
- Exploding information volume
- Distributed decision environments
- Loss of institutional knowledge.

The number and complexity of repetitive tasks is increasing. In today's automated environment, the information needed to make a decision is available, but only if the knowledge worker knows WHERE and HOW to access the information. Its typically located on a variety of computers, each with its unique operating system, applications packages, query structures and file formats. Remembering the steps to access information across these environments is difficult and error prone. In the past, clerical staff performed repetitive tasks such as updating letters and documents. Now, knowledge workers do this with computers and wordprocessing systems.

Information is exploding in volume. Yet effective decision making demands the ability to correlate and evaluate information quickly and rapidly. The quantity of information is rising at an exponential rate

(doubling every 10 years), but the ability to cope with this growing mass of information is rising linearly, at best. In addition, the complexity of organizations is increasing, especially in terms of interrelationship of functions. No longer can we classify decisions as "comptroller" or "personnel office." Knowledge workers must make decisions across traditional functional areas. This requirement necessitates access to information outside the traditional stovepipe information systems.

Organizational and decision environments are increasingly complex. The decision making process now requires more players, more information and more communication. Knowledge workers face both a cyclic set of tasks but also an increasing number of ad hoc tasks. These tasks, because of their high visibility and short term nature, tend to receive top priority. The longer range and once-a-year tasks fade into the background until their due dates are imminent. Shifting events and dates produce complex and dynamic schedules. The resultant changes are often "lost in the mail" and do not reach the knowledge worker until it is too late for an effective response.

Turnover of personnel is always an issue and a complex decision environment exacerbates the problem. A new knowledge worker may understand neither the process nor the product. A motivated knowledge worker will seek guidance from others, which causes a loss in productivity for the knowledge worker's tutor. Attempts to retain institutional knowledge through regulations and procedures are effective only when tasks are very specific and well defined. Higher order decisions are a value added function and not amenable to written procedures. In addition, keeping procedures current is difficult and expensive.

The Objective

The objective of the Knowledge Worker System (KWS) is to provide a synergistic performance support environment that assists Army knowledge workers in fulfilling their responsibilities. KWS is not an application designed to perform a specific function but rather an integrated environment that interrelates existing applications and multiplies knowledge worker productivity within a complex task environment.

KWS will remind personnel WHAT tasks to perform and WHEN; the system will retain procedural information (HOW to accomplish) and automate many repetitive tasks. By enhancing productivity, KWS will free up personnel to do the creative and complex tasks in which humans excel.

Productivity savings in the Army infrastructure translate to additional resources available for the combat soldier. We estimate

that implementation within just the Corps of Engineers will result in an annual savings of \$120 million.

The Knowledge Worker System will provide the following capabilities:

- Perform dynamic scheduling for organizational tasks.
- Capture and store institutional knowledge.
- Provide procedural information about what and when a task must be done and how it was historically performed.
- Provide a framework for attaching expert systems and other productivity tools.

Enable Dynamic Scheduling

Knowledge workers' daily activities are dictated by pre-determined tasks that are part of a prescribed sequence of events, which lead to a completed process. Generally, higher headquarters determines the overall process and calendar; thus, the knowledge worker has little control of timing and due dates. The knowledge worker initiates actions, analyzes information and produces products according to the preset schedule.

This schedule, albeit preset, is subject to change. Changing the schedule is easy; determining the effect and notifying the critical knowledge workers is much more difficult. Task completion requires careful work planning, scheduling and coordination with related activities and a myriad of ad hoc tasks and meetings. Further, as the task progresses, it must be periodically re-planned, re-scheduled and re-coordinated. KWS will accommodate changes in the work schedule by reflecting their effects on related activities, adjusting due dates and providing information about the impact of the change.

The Dynamic Scheduler will provide both a multi-user calendar and the ability to move tasks within boundaries and determine the effects of proposed changes. The multi-user calendar will consist of four tiers:

- Level 1: Major, headquarters milestones
- Level 2: Organizational milestones
- Level 3: Individual milestones
- Level 4: Personal dates

Milestones set at higher levels link to lower level tasks and therefore, the effect of proposed changes can be appropriately reflected. For example, a relatively small change in a level 1 milestone may unacceptably shorten the time available for a task

at level 3. Figure 1 shows the cascading work hierarchy. Each event drives an organizational responsibility and has a due date. Events break down into tasks, for which durations can be specified. Tasks can be further decomposed into subtasks. All this information can be considered as the WHAT and WHEN drivers for knowledge workers.

Steps are the procedural information that specifies a task's historical performance. Regulations, examples from previous years and other reference materials make up supporting documents that supplement this procedural information. This information can be considered the HOW drivers for knowledge workers.

Capture and Store Institutional Knowledge

The capture and storage of institutional knowledge is critical to the success of KWS. The initial information regarding tasks, processes, steps and calendar information can be gathered and entered a variety of ways. A person who is familiar with the overall function or mission can observe job practices and capture those tasks and procedures. Each task is then assigned to a particular knowledge worker. Computer-Assisted Software Engineering (CASE) tools also provide a mechanism to decompose processes into tasks and information flow. The difficult part is to capture task steps and products. This decomposition should be done as a natural by-product of the task performance, not as an added function.

Each task, and its related steps, have an informational and procedural content. The information for each step will be attached to the step. Thus, the knowledge worker can update existing information rather than start anew. The information also

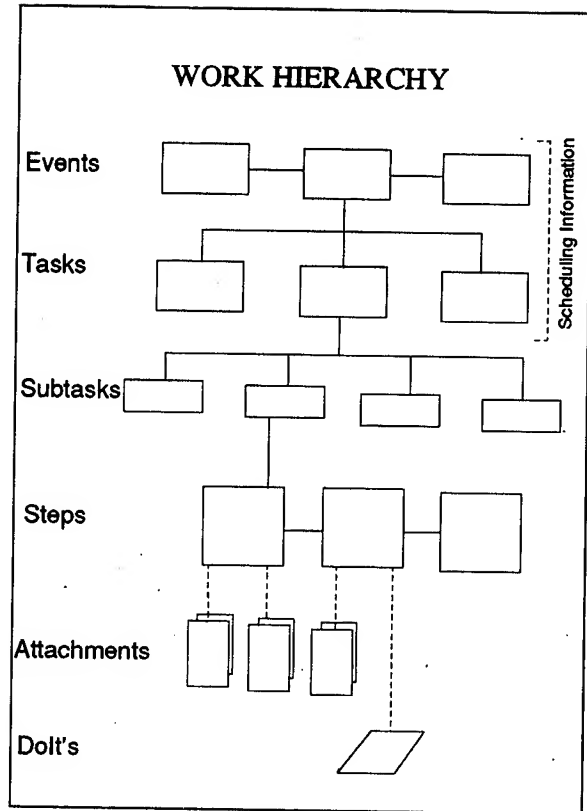


Figure 1

KWS is structured around a hierarchy of events. At the highest level are the events that comprise the organizational function. Events are decomposed into tasks, subtasks and finally into steps for accomplishment.

must be characterized as to subject, task and knowledge worker so that it can be located and used as a template for new tasks.

The present system of "paper storage" is not adequate to support a knowledge worker. Too often the paper record cannot be located. In addition, the document provides only the result, with no hint about the decision process, the base data, or the location of automated records. Simply saving data in machine readable form is not adequate. Finding the "right data" among the 200 diskettes left by a departing employee is a futile task.

Provide Procedural Information

The procedural information about what, when and how a task was historically performed is a key element in KWS. The Dynamic Scheduler and Calendar provide the initial information, with knowledge worker input of ad hoc tasks. KWS will be able to check the organization's calendar, update the individual calendar and notify the knowledge worker of conflicts among tasks.

The dynamic scheduler furnishes the information central to the organization's goals and process. It provides a description of the task along with the steps required for performance. Each step can contain examples of documents and forms, references to regulations, historical notes and links to related steps in other tasks. The knowledge worker can easily modify the steps to incorporate new information. This process not only helps the current knowledge worker keep track of current progress but retains institutional knowledge. Humans tend to forget the details of tasks and when called on to repeat, must spend considerable time answering the question "what did I do last time?"

Steps can be constructed in a variety of ways. Initially, knowledge acquisition may be accomplished by either an interview process or via a knowledge worker's direct entry. The latter is time-consuming but has the useful by-product of being a place for the knowledge worker to collect thoughts and ideas before tackling the task. In the future, step information will be dynamically captured, either directly during task performance or by use of "step templates" from a task pool. The step information allows a new knowledge worker to benefit from others past experience.

Additionally, many complex organizations have field units, each of which has the same basic set of operating procedures. Once a Knowledge Worker environment is established and the tasks/steps information captured; it can be mapped to a new location with only minor changes. Minor variations in the steps

may require some customization but the basic step information, regulations, help information and products will be similar.

Provide a Framework for Attaching Productivity Tools

The knowledge worker performance support environment provides a framework for attaching applications. The environment still requires conventional applications: word processing, spreadsheets, data base packages and specialized function, but will help determine needed tools and when to apply them. The knowledge worker must perform the value added functions, particularly for the ill defined tasks.

KWS invokes conventional applications from related steps. When the knowledge worker selects a step, the environment knows the type of information to be manipulated and invokes the appropriate tool. Thus, if the step is to modify a letter, the environment will invoke the word processor.

Many tasks that clutter up the work day will be automated with "DoIts." Tasks such as accessing a mainframe and downloading data, initiating batch processes or transferring files to other knowledge workers can be entirely automated. Steps that have this type of automation will have a DoIt option. Selecting this option will initiate the process and allow the knowledge worker to proceed with other activities.

The Approach

Knowledge Worker System and other performance support systems can be classified as groupware, i.e., software that facilitates teamwork across time and space. Groupware provides the means for a team of people to work on a project concurrently or independently. Other aliases for this software genre include "computer-supported cooperative work", "collaborative computing" and "computer-assisted communication."

Early in the development process, we realized that a well-designed user interface as a critical component for KWS. In order to provide a system that assists users throughout their working day, the interface between the user and the system capabilities should be functional, intuitive and attractive. This same requirement can be applied to groupware as a class of software.

System developers also recognized the futility of designing an innovative system without the involvement of its targeted users. Initial attempts to involve users indicated two obstacles:

- Users had no frame of reference for this type of system and had difficulty visualizing the concept
- Users had limited time to devote to the system design process.

Therefore, we selected a process of iterative refinement as the system development approach. See figure 2.

We selected Windows by Microsoft and Actor by The Whitewater Group as the initial development platform. Windows provides a multi-tasking environment and a graphical user interface. Actor is an object-oriented programming tool set that expedites the development process and simplifies iterative prototyping.

USACERL has produced several prototypes that model KWS features to evoke user involvement in refining system functionality. We used an IBM Joint Application Design (JAD) methodology that employs structured, facilitated, user-centered workshops to solicit user feedback and to generate user requirements for the system. We have emphasized the evaluation of the user interface in this process.

Enhanced prototypes that function as limited working versions of the system will be placed in user offices as pilot systems. Evaluation of these pilots will provide refinements to the system requirements and spawn other production versions. These pilots will be working versions of the system and therefore also will provide service to the pilot sites.

After the system development effort stabilizes, KWS will be ported to a delivery platform, e.g., OS/2 or Xenix. We made the initial choice of DOS/Windows to shorten the development cycle; however, multiple hardware/software environments and performance issues drive the choice of delivery platform.

An additional benefit of delaying the delivery platform decision is that hardware and software continue to evolve as we refine KWS's functionality. Additional platform choices, that were under

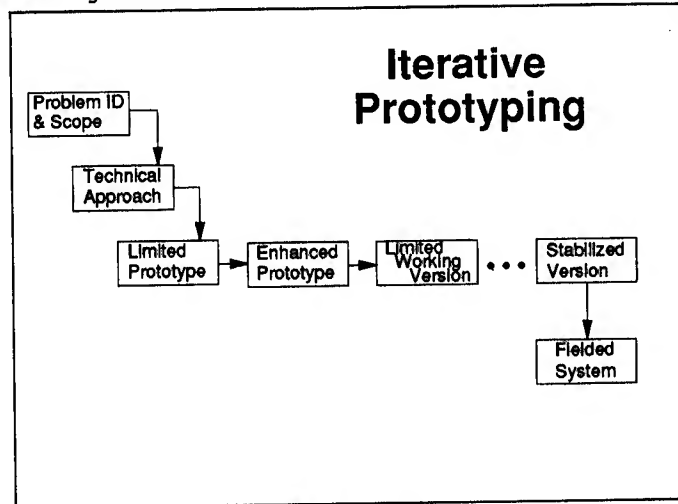


Figure 2

Iterative prototyping was used to maximize user involvement and feedback during the development process.

development at the initiation of KWS, are likely to be viable options in the deployment phase.

Status

Knowledge Worker System (KWS) addresses the need for a highly intelligent, innovative system that will enhance productivity, capture institutional memory and reduce the stress on personnel that is prevalent in an event-driven, high-demand environment. The Corps of Engineers' Research and Development Directorate tasked USACERL to perform research and system development. Mr. John Sheehey III, Corps of Engineers Headquarters, is the technical monitor and research sponsor. Mr. Sheehey envisions a system that will make the decision making process in the Military Construction, Army (MCA) cycle more timely, predictable and less stressful.

The Knowledge Worker System effort began in 1987 as a project to develop a standard operating procedure (SOP) handbook for the position of the chief of the Budget Book production process. In the first quarter of 1988, USACERL compiled the lessons learned from the SOP handbook efforts, and the collective efforts of thirteen researchers into a white paper that described KWS as "an attempt to capitalize upon the benefits of automation in order to serve humans in any organizations that rely upon automation systems to support their activities." In retrospect, we would classify this system as a performance support environment.

In the following quarter, we launched a project to analyze the connectivity between the six offices that were the key players in the MCA budget process. We used a Computer-Assisted Software Engineering (CASE) tool to store the resulting information because of the tool's ability to update globally all references to a given piece of data and its ability to display graphically relationships between data. KnowledgeWare's Information Engineering Workbench is the CASE tool used for the work.

At the conclusion of this project, we asked the participants to take part in a User Group to design the user interface and KWS's functionality. The first effort was participation in a structured workshop to validate the prototype and collect system requirements for future versions of KWS.

The development strategy included a prototype system because of three key realizations:

- The need to have potential system users closely involved in the system design process;

- The efficacy of evoking user involvement by providing a subset of system functions;
- The cost-effectiveness of developing an innovative system in an iterative progression.

We contracted with Georgia Institute of Technology (GA Tech) to design and develop the KWS's prototype. In Fall, 1989, we installed the KWS prototype on four workstations MCA Budget Office. We demonstrated the prototype to the User Group and promoted it as a vehicle for eliciting user response. This subset of KWS functions has the look and feel of a production system while avoiding the costly and lengthy development cycle of a fully functioning system.

In November 1989, the KWS User Group spent three days at the Fort Belvoir Fusion Center in a structured workshop that focused on generating requirements for the next iteration of KWS. Both users and developers responded positively to the workshop and advocated future workshops to expand upon system requirements. The iterative system development concept is a necessary component for Knowledge Worker System's success.

The next step is to refine the system specifications generated by users/developers at the November 1989 JAD Workshop and initiate development of the KWS pilot system. We are reviewing the software platform used for the KWS prototype, i.e., Windows and the Actor programming language. While this choice proved cost-effective as a prototype platform, it appears that this platform will be insufficient to meet user requirements in the long term.

We are analyzing the relative advantages/disadvantages of a) modifying the prototype at this point versus b) changing platforms in order to satisfy long term system requirements. A significant aspect in this analysis is the realization that changing platforms will result in a longer development period before the pilot version of the system (KWS, Version 1.0) is available. Conversely, retaining the current platform will mandate a major overhaul when current capabilities prove inadequate to fulfill user-defined system requirements.

Following this pro/con analysis, we will make recommendations to the system proponent and users. By the beginning of the second quarter FY 90, we will provide the KWS User Group with recommendations for the next iteration of the system, KWS Version 1.0. The User Group will be invited to participate in another structured workshop with primarily focus on scheduling issues.

We will produce two development plans: one addressing the fielding of KWS, Version 1.0; the second describing KWS development through FY 91. Both plans will be available in early Spring, 1990.

Image Processing Applications Via Photorefraction (U)

***Edward J. Sharp, Dr., William W. Clark, III, Dr.,
Mary J. Miller, Mrs., and Gary L. Wood, Mr.
CECOM Center for Night Vision and Electro-Optics
Fort Belvoir, Virginia 22060-5677**

INTRODUCTION

A large optical nonlinearity can be obtained in photorefractive materials with low-power cw lasers. This feature, combined with the fact that the nonlinearity in photorefractive materials is a nonlocal response, has led to the demonstration of a wide variety of prototype devices for real-time holography and optical processing applications.¹ The motivation for much of this research has been provided by the realization that optical processing offers several unique advantages for high-speed information processing. Consider, for example, the use of the mathematical operation of two-dimensional cross-correlation for determining the degree of similarity between two optical images. This operation has been identified as essential for a number of commercial and military applications including robotic vision, spacecraft docking, automatic target recognition, and tracking. Even for images of modest resolution, such as those produced by television imaging devices, the calculation of the cross-correlation function is very extensive exercise in computation. This difficulty is compounded when the computations must be made in "real-time" (approximately 30 sec⁻¹ or a TV-frame rate) since the number of mathematical operations such as additions or multiplications can be on the order of 10¹⁰. It has been recognized that this level of performance can only be achieved through coherent optical image processing.

When the information to be processed is in image form, photorefractive approaches for tasks such as image amplification, volume storage, addition/subtraction, restoration (aberration correction), and edge enhancement are

very attractive. Optical systems which carry out these image manipulations require the important step of transferring the real-world scene contained in an incoherent light beam to a coherent laser beam. That is, a spatial light modulator (SLM) is required). Ultrafast speeds are not required for many of these applications because ultrahigh bit rates can be achieved by performing numerous nonlinear processing operations in parallel over two-dimensional fields. Two-dimensional SLM technology is an area of intense interest with over fifty approaches being developed.² Candidate SLM devices based on magneto-optic, electro-optic, deformable membranes, photochromic, thermoplastic, photorefractive, liquid crystals, multiple quantum wells, ferroelectric liquid crystals and organic optical materials have all been the subject of considerable study and development. Incoherent-to-coherent conversion has been successfully demonstrated by a number of researchers using photorefractive crystals as SLMs.^{3,4}

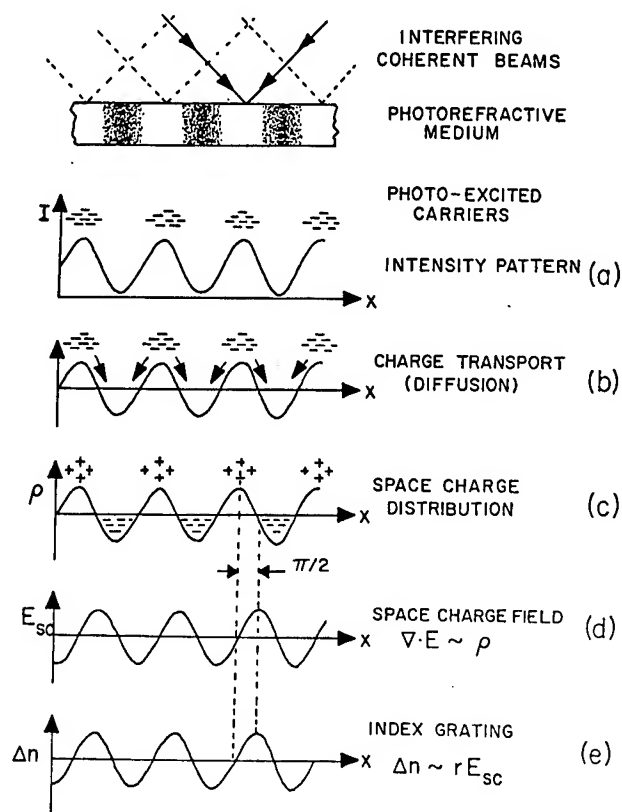
In addition to SLMs a host of other important image processing applications have been brought forth through the utilization of the photorefractive effect including real-time holography, image amplification, correlation-convolution operations, edge enhancement, motion detection and tracking, phase conjugation for image restoration and beam clean-up, incoherent-to-coherent light conversion, beam steering, reconfigurable interconnections, real time image subtraction, associative holographic memories, optical switching, and bistability. These optical image processing applications have already been demonstrated and there are new photorefractive effects and materials still being discovered which will certainly lead to the demonstration of even more applications.

We shall briefly describe the details of the photorefractive mechanism and present recent experimental results which demonstrate the utility of photorefractive materials for real-time volume storage of color images, image addition/subtraction, and image restoration (aberration correction).

PHOTOREFRACTIVE MECHANISM

The photorefractive optical nonlinearity (light induced refractive index change) is one of the largest observed ($n \sim 10^{-3}$). There are several differences which set the photorefractive nonlinearity apart from other light induced

Figure 1. The formation of photorefractive index gratings. (a) Periodic light intensity pattern and photo-excited charge carriers. (b) Charge transport by diffusion. (c) Static charge distribution. (d) $\pi/2$ phase-shifted space charge distribution. (e) Photorefractive index grating.



index changes. Briefly, the magnitude of the steady-state photorefractive index change is independent of the total intensity of the incident beams and is spatially non-local (the maximum refractive index change does not necessarily occur where the beam intensities are greatest). However, the time for formation does depend on the beam intensity.

The microscopic details of the photorefractive mechanism is normally described using a band transport model which assumes the existence of a pool of charges residing in low-lying traps. The physical basis of the photorefractive mechanism can be understood by examining Fig. 1. In the presence of nonuniform or spatially periodic illumination (Fig. 1a) the charges (either electrons or holes) are excited from defect sites or impurity centers leaving behind ionized donor sites. The photo-excited charges migrate by diffusion (Fig. 1b) out of the illuminated regions and are eventually retrapped in the dark regions of the crystal. A static charge distribution (Fig. 1c) is formed in the crystal due to

the spatial separation of the ionized donor sites in the illuminated regions and the retrapped charges in the dark regions. This space-charge distribution is balanced by a strong space-charge field according to Poisson's equation. The space-charge-field distribution (Fig. 1d) is shifted spatially ($\pi/2$ degrees) relative to the incident intensity distribution. This strong electrostatic field ($\sim 10^5$ V/m) then produces a change in the refractive indices (Fig. 1e) through the linear electro-optic effect. If the incident intensity pattern is periodic (i.e. the intensity pattern is produced by interfering two coherent laser beams) then the resulting index change will also be periodic, resulting in the formation of an "index grating". The $\pi/2$ phase shift between the incident intensity pattern and the index grating causes the two interfering or "writing" beams to couple as they propagate through the crystal. In this two-wave mixing process one of the beams experiences gain (amplified) and the other loss (deamplified).⁵ Feinberg et al. demonstrated that one beam can deplete essentially all of the power of the other beam after propagating through only a few millimeters of a BaTiO_3 crystal.⁶

PHOTOREFRACTIVE EFFECTS

To obtain a better understanding of the prototype operations demonstrated in photorefractive materials we briefly review the photorefractive effects which lead to these applications. As discussed earlier, two-beam coupling results in energy exchange between the interfering beams and this phenomena is responsible in turn for a multitude of other photorefractive effects. In materials with high gain such as BaTiO_3 , SBN, and BSKNN it is easy to observe beam fanning (asymmetric beam scattering)⁷, phase conjugation,^{8,9} (degenerate four-wave mixing), double-pumped phase conjugation,^{10,11} and cooperative (multicolor) effects.¹²

Beam Fanning

When a single laser beam is passed through a photorefractive crystal, energy exchange or two-beam coupling takes place between beams comprised of scattered light and the incident beam itself. This effect called "beam fanning" forms the basis for many of the interesting phenomena observed in photorefractive materials. Beam fanning can be briefly described as follows: Scattered coherent light from crystal imperfections and impurities (initially very weak in

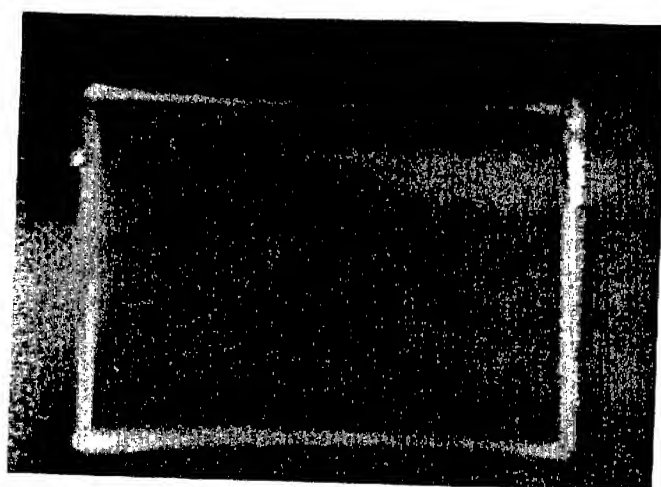
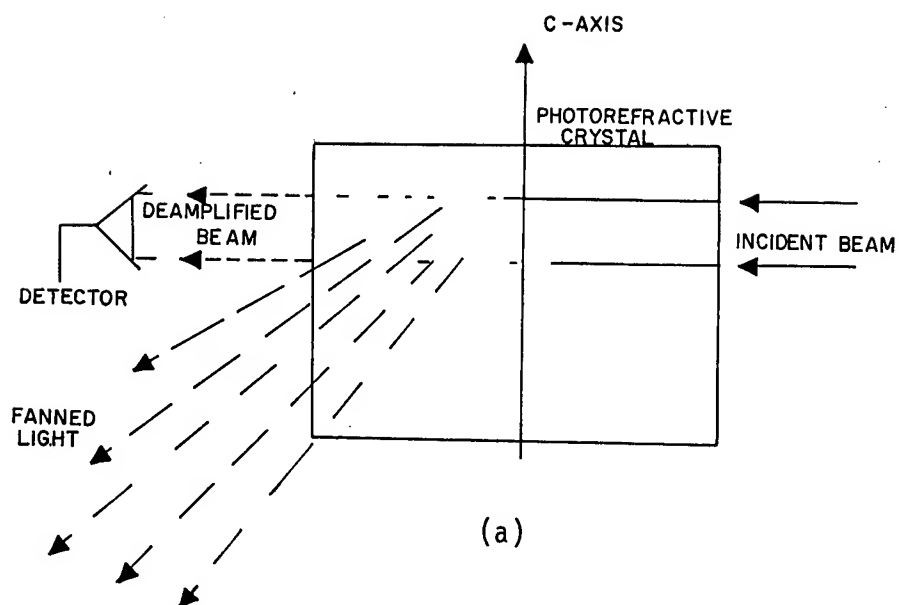


Figure 2. (a) Illustration of photorefractive beam fanning based on the photograph of Fig. 2b. Scattered light is amplified resulting in a broad fan of light and a deamplification of the incident beam. (b) Microphotograph (X 10) of a Ce-SBN:60 crystal while fanning a beam incident at 90° to the c-axis (directed from bottom to top).

intensity) crosses some portion of the input beam and two-beam coupling occurs. The scattered light propagating in the directions providing gain will be amplified at the expense of the input beam. As the scattered light builds in intensity additional gratings are formed which weakens the transmission in the direction of propagation and produces a broad fan of light to one side of the main beam as depicted in Fig. 2. Figure 2a is a sketch of a microphotograph shown in Fig. 2b. The crystal used is Ce-SBN:60 and the extraordinary polarized beam is incident on the crystal normal to the c-axis. The input beam depletion can be significant (transmission ~ 0.1%).⁹ The buildup time of the scattered light does not follow the two-beam coupling response time and the statistics of the scattered light is Poissonian. The amount of scattered radiation will depend on the magnitude of the photorefractive gain coefficient, and the effective interaction length. The direction of energy exchange will depend on the electro-optic tensor elements and their relative magnitude. The amount of steady-state beam depletion and the time to reach depletion (time response) are the essential parameters of the beam fanning process from an applications point-of-view.

The effect of beam fanning is not reserved for single wavelength sources. In fact, for particular beam fanning geometries many mutually incoherent (different wavelengths) beams can produce an intense beam fan. When a multicolored laser source such as an argon-ion laser, simultaneously oscillating at eight different wavelengths, is passed through a photorefractive crystal a multicolored beam fan is produced.¹² The basis of this phenomena can be explained by examining Fig. 3a and 3b. Incident light of one wavelength (blue for example) along direction $(\mathbf{k}^i)_B$ produces a fan of scattered light with different values of grating wavelength or periodicity indicated by different values of the grating wave vector $(\mathbf{k}^g)_B$. When more than one color is present in the incident beam, both produce their own fan patterns. However, at one particular angle (for each angle of incidence), as shown in Fig. 3b, the grating wavelength for incident blue light, for example, will be the same as for the slightly dispersed incident green light. At this angle the blue light is Bragg matched to the grating written by the green light and vice versa, i.e., $(\mathbf{k}^g)_B = (\mathbf{k}^g)_G$. This in turn strengthens the grating (i. e., a larger grating modulation ratio), and more scattering occurs which

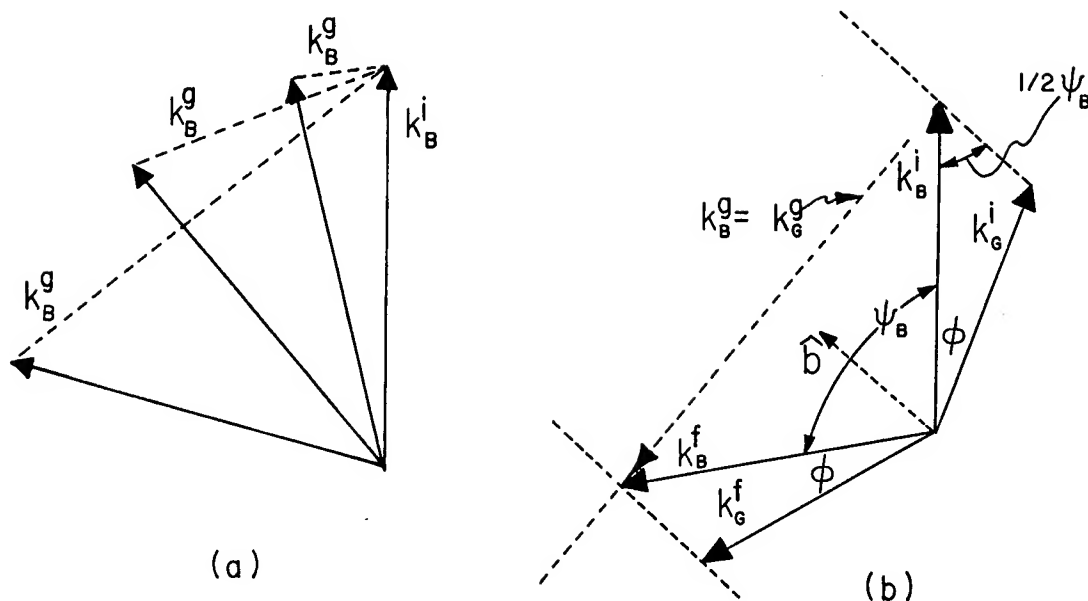


Figure 3. (a) Resulting grating wave vectors for fanned light produced by a single color (blue) input wave k_B^i . (b) Wave vectors for two different colors (blue and green) as seen in the plane of incidence. Although beam fanning produces fanned light in many directions, only at one selective angle are the grating wave vectors equal in both direction and magnitude.

strengthens the grating even more. This bootstrapping process continues until this one grating dominates all other gratings. This mechanism therefore allows the coupling of mutually incoherent beams and an intense multicolor cone or "rainbow" of light emerges from the crystal in a time nearly equal to the beam fanning response time.¹²

Phase Conjugation

Phase conjugation via four-wave mixing geometries are easily achieved using photorefractive materials with cw lasers having only milliwatt incident intensities.¹³ The two principle ways to generate phase conjugation in photorefractors via four-wave mixing has been; 1) to use external mirrors in a number of different configurations, and 2) to use internal reflections within the crystal to generate the required beams. The later method results in the so called self-pumped phase-conjugate mirror (SPPCM).⁸ In order

to observe SPPCM the gain-length product needs to be above some threshold value. A two-interaction region model has been proposed to describe the SPPCM which assumes that radiation scattered out of the input beam is retroreflected from a corner of the crystal back onto itself.¹⁴ Two regions of four-wave mixing are established along the input beam which are coupled together by these retroreflected counterpropagating beams. Both interaction regions contribute to the phase conjugate output. Figure 4 shows a microphotograph of a crystal of Ce-doped SBN:60 while it is producing a self-pumped phase conjugate wave.¹⁵ The loop formed in the corner of the crystal is typical of the tungsten bronze family of photorefractive crystals and contains the counterpropagating beam required for the two-interaction region model.

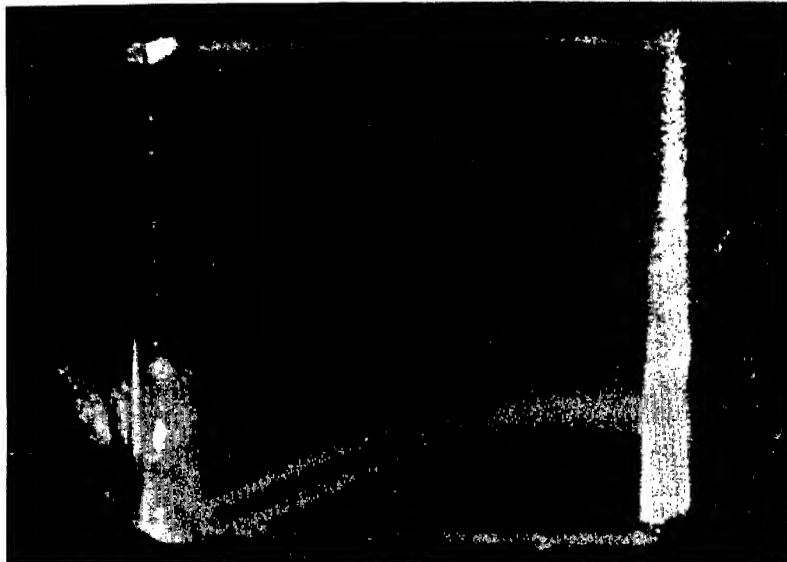


Figure 4. A microphotograph of a Ce-SBN:60 crystal showing the characteristic corner loop observed when the crystal is producing a conjugate wave via self-pumped phase conjugation. The input beam is extraordinary polarized and enters the crystal from the right at normal incidence. The c-axis of the crystal is directed from bottom to top.

The buildup time of the conjugate reflectivity from a self-pumped phase conjugator is like the buildup time of a

laser. The initial phase conjugate beam builds up from noise. In fact, since the beams are often observed to "bend" from the input beam into a crystal corner, the role of secondary gratings is important. These secondary gratings are written initially by beams of low intensity and the dark contribution should be important. As a result the gain of these gratings will be intensity dependent. The important parameters which describe four-wave mixing and which are important from an applications point-of-view are the phase-conjugate reflectivity and the buildup time (initiation time) of the phase conjugate wave.¹⁶

Double Phase Conjugation

One of the more recent developments in the area of phase conjugation has been the demonstration of several different geometries in photorefractive media for the construction of double phase conjugate mirrors (DPCMs).^{11,17} In these devices the photorefractive material acts as a conjugate mirror for two incoming waves simultaneously. DPCMs require no external pumping beams and the two incident beams need not be coherent with each other. An example of this device is depicted schematically in Fig. 5. Here the arrangement used

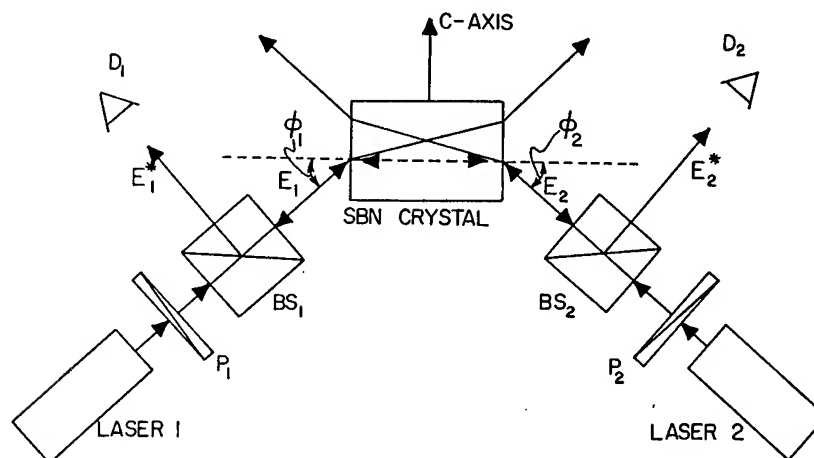


Figure 5. Experimental apparatus used to study the reflectivity, angular sensitivity, and time response of the bridge conjugator. D-detector, P-polarization rotator, and BS-beam splitter.

is called the bridge conjugator because the two mutually incoherent beams (different lasers with nominally the same wavelength) enter opposing [100] SBN:60 crystal faces and fan together to form a bridge without reflection from a crystal face. A phase conjugate replica of each beam appears and both beams are required for either conjugate to exist. The energy for the conjugate beam being supplied by the other beam and vice versa. For example, if beam 1 is blocked the conjugate of beam 2 disappears instantly while the conjugate of beam 1 disappears with the characteristic erasure time of the crystal. A similar behavior is observed if beam 2 is blocked. In this conjugator beam 1 writes the gratings responsible for the production of its conjugate while beam 2 reads these gratings and supplies the energy for the conjugate. This behavior is a characteristic of all DPCMs and makes possible several image processing applications.

APPLICATIONS

Real-Time Volume Storage of Color Images

We pointed out above that double phase conjugate mirrors do not require that the beams be coherent with each other. They do, however, need to be of the same wavelength (to satisfy the Bragg condition) or beam steering will result.¹⁸ Early attempts to produce full-color phase-conjugate replicas¹⁹ met with limited success due to grating competition and crosstalk between the different laser wavelengths. The effect of multicolor beam fanning has been used to overcome the difficulty of erasure between colors when a multiwavelength source is used to produce phase conjugate waves.²⁰

As shown in the experimental arrangement of Fig. 6, the multicolor rainbow pattern produced by a cerium doped SBN:60 crystal was back reflected into the crystal giving rise to four-wave mixing and phase conjugation. Color phase conjugation was obtained without competition and crosstalk between colors. Color images of a centimeter in size were stored in the crystals of SBN:60, SBN:75, BSKNN II, and BSKNN III with a resolution exceeding 30 line pairs per millimeter. The phase conjugate images were produced with a reflectivity of 10%.

Image Amplification and Aberration Correction

Two image processing applications, amplification and aberration correction, have been demonstrated by many workers

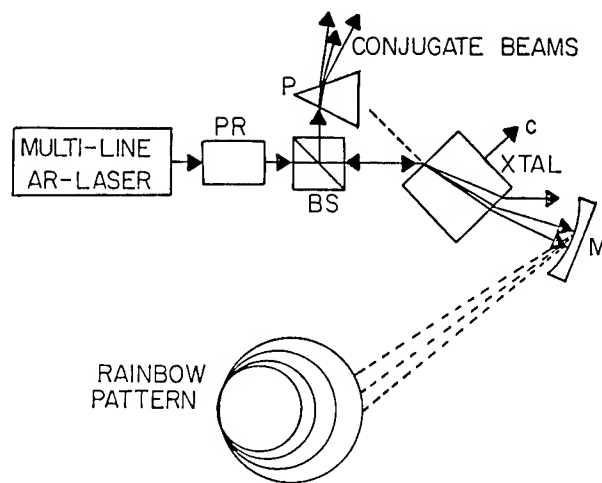


Figure 6. The rainbow pattern is back reflected by mirror M to produce a multicolored conjugate wave. An illuminated transparency was used to generate the colored phase-conjugate images. The individual colored images were examined by using wavelength selective filters instead of a dispersing prism.

in a number of different photorefractive materials. Coherent image amplification using cw lasers and photorefractive crystals has been carried out in two-beam coupling and four wave mixing arrangements. Two-wave mixing is a very powerful technique for achieving beam amplification.²¹ A coherent image amplifier utilizing two-wave mixing in BaTiO_3 was recently analyzed to determine the conditions necessary for optimal performance. Experimental results on this amplifier verified that a gain of 4000, grey-level dynamic range greater than 100, and a resolution of 250lp/mm could be achieved, giving rise to a space-bandwidth product of 10^6 for a (4 mm X 4 mm) aperture BaTiO_3 crystal.²²

When two-wave mixing is used to amplify images severe phase aberrations may be introduced by the photorefractive gain medium unless the crystal surfaces are prepared very carefully and the medium is free from all index variations. One technique to avoid this potential problem is to use a self-pumped phase conjugate mirror (SPPCM) for the restoration of the phase aberrated laser beam. The SPPCM described earlier is ideally suited for a number of beam clean-up and image restoration applications since only one

input beam is required to generate the conjugate signal. These devices offer the advantages that they do not require an external laser to provide high quality counterpropagating pump beams and they can produce conjugate signals for very low power light levels.

Using the experimental set-up depicted in Fig. 7, and a single crystal of SBN:60 we demonstrated the distortion correction (beam clean-up) property of the conjugate wave.

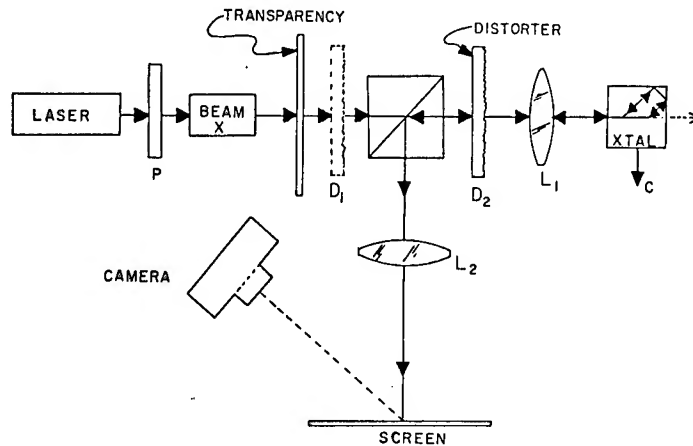


Figure 7. Apparatus used to produce phase-conjugate replicas of input images and to test the aberration correction ability of a self-pumped phase conjugate mirror. Beam X-beam expander.

Figure 8a shows the input scene located in the object plane (transparency) which is imaged by lens L_1 into the crystal. The reflected beam is then re-imaged by L_1 and L_2 into the image plane (screen) and displayed in Fig. 8b. As can be seen, the image is a faithful reproduction of the "Conquest of Darkness" emblem. In fact, the image is the emblem's phase conjugate. This is clearly demonstrated by adding a wavefront distortion to the input beam at D_1 . The phase-conjugate reflected beam examined at the screen does not retrace its path through the aberrator, and hence, cannot become undistorted as shown in Fig. 8c. However, when the wavefront distorter is placed at D_2 the phase-conjugated reflected beam does retrace its path through the distorter and becomes undistorted as seen in Fig. 8d. The reconstruction of the object after a double pass through a distorter

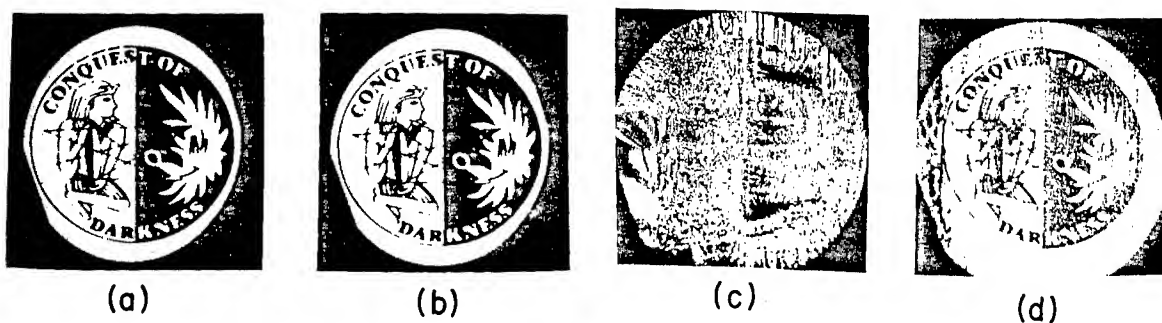


Figure 8. Photographs demonstrating phase aberration correction. (a) Input image. (b) Phase-conjugate replica of input image. (c) Phase-conjugate replica of phase-aberrated input image. (d) Phase-conjugate replica of input image after correction.

verifies that the image of the emblem seen in Fig. 8d is indeed a phase conjugate.

Addition/Subtraction

Several methods of real time optical image subtraction have been proposed which take advantage of photorefractive media. We have performed image addition/subtraction using a single crystal of SBN:60 in a DPCM arrangement.¹¹ The experimental arrangement in Fig.9 shows phase conjugate

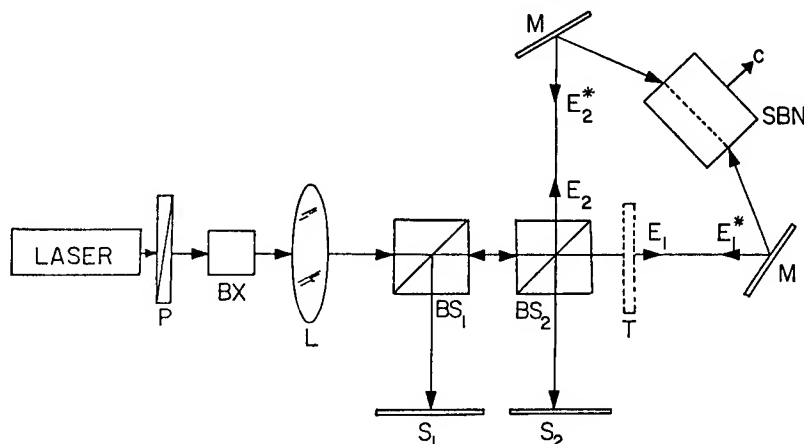


Figure 9. Experimental apparatus used to observe addition and subtraction of images and the fidelity of phase conjugation for the bridge conjugator.

waves E_1^* and E_2^* produced in the SBN crystal being recombined at BS_1 and BS_2 . In this experiment the two holograms containing image information with an inherent phase shift of 180° (obtained by virtue of the Stokes principle and beam mixing on a beam splitter) were recorded simultaneously in the crystal. E_2^* has a near constant intensity while E_1^* is the conjugate of the input image (transparency). The conjugate image of the object slide produced at S_1 (shown in Fig. 10a) is an exact replica of the input image formed from the sum of E_1^* and E_2^* due to constructive interference (addition) at BS_2 . A 180° phase shift between E_1^* and E_2^* is introduced by BS_2 and the conjugate image at S_2 shown in Fig. 10b is an exact negative of the object slide as a result of destructive interference (subtraction) between the conjugate waves E_1^* and E_2^* . These results are unchanged when a constant phase shift is placed in the path of E_2 or when the optical path of beam E_2 is physically changed. Methods of image addition and subtraction which utilize DPCMs appear to be less sensitive to alignment and intensity variations compared to phase conjugate interferometric techniques.

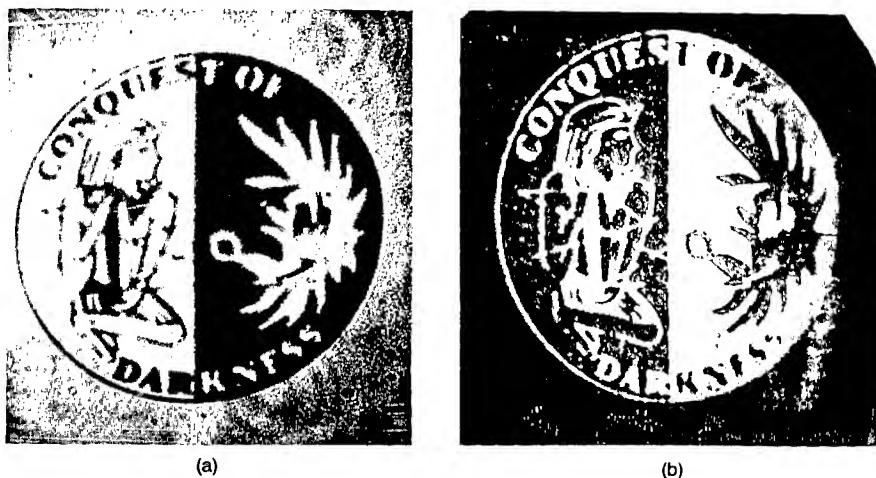


Figure 10. (a) Photograph of E_2^* added to E_1^* taken at S_1 in Fig. 9. (b) Photograph of E_1^* subtracted from E_2^* taken at S_2 in Fig. 9.

REFERENCES

1. P. Gunter, Phys. Rept., **93**:199-299, (1982)
2. A. D. Fisher and J. N. Lee, SPIE Proc., **634**:352, (1986).
3. B. A. Horwitz and F. J. Corbett, Opt. Eng., **17**:353, (1978)

4. Y. Shi, D. Psaltis, A. Marrakchi, and A. R. Tanguay, Jr., Appl. Opt. **22**:3665, (1983)
5. G. L. Wood, W. W. Clark, III, M. J. Miller, E. J. Sharp, G. J. Salamo, and R. R. Neurgaonkar, IEEE J. Quantum Electron., **QE-23**:2126 (1987).
6. J. Feinberg, D. Heiman, A. R. Tanguay, Jr., and R. W. Hellwarth, J. Appl. Phys., **51**:1297 (1980)
7. J. Feinberg, J. Opt. Soc. Am. **72**:46 (1982)
8. J. Feinberg, Opt. Lett. **7**:486 (1982)
9. G. J. Salamo, M. J. Miller, W. W. Clark, III, G. L. Wood, and E. J. Sharp, Opt. Commun., **59**:417 (1986).
10. S. Weiss, S. Sternklar, B. Fischer, Opt. Lett. **12**:114 (1987)
11. E. J. Sharp, W. W. Clark, III, M. J. Miller, G. L. Wood, B. Monson, G. J. Salamo, and R. R. Neurgaonkar, Appl. Opt., **29**:743, (1990).
12. G. J. Salamo, M. J. Miller, W. W. Clark, III, G. L. Wood, E. J. Sharp, and R. R. Neurgaonkar, Appl. Opt., **27**:4356, (1988).
13. J. Feinberg, "Optical Phase Conjugation in Photorefractive Materials", in Optical Phase Conjugation, R. A. Fisher, Ed., New York, Academic Press (1983)
14. K. R. MacDonald and J. Feinberg, J. Opt. Soc. Am., **73**:548 (1983)
15. J. Rodriguez, A. Siahmakoun, G. J. Salamo, M. J. Miller, W. W. Clark, III, G. L. Wood, E. J. Sharp, and R. R. Neurgaonkar, Appl. Opt., **26**:1732 (1987)
16. M. J. Miller, E. J. Sharp, G. L. Wood, W. W. Clark, III, G. J. Salamo, and R. R. Neurgaonkar, Opt. Lett., **12**:340, (1987).
17. P. Yeh, T. Y. Chang, and M. D. Ewbank, J. Opt. Soc. Am. B **5**:1743-1749, (1988)
18. B. Fischer and S. Sternklar, Appl. Phys. Lett., **51**:74 (1987)
19. M. Cronin-Golomb, S. Kwong, and A. Yariv, Appl. Phys. Lett., **B4**:727 (1984)
20. G. J. Salamo, B. Monson, E. J. Sharp, W. W. Clark, III, G. L. Wood, M. J. Miller, and R. R. Neurgaonkar, Electro-Optics and Nonlinear Optics: Symposium VIII, Ceramic Science and Technology Congress, Anaheim CA., (1989) Paper 25-SVIII-89C
21. A. Marrackchi and J. P. Huighard, Appl. Phys., **24**:131 (1981)
22. Y. Fainman, E. Klancnik, and S. H. Lee, Opt. Eng., **25**:228 (1986)

SHEPARD AND SASS

Thermal Imaging at Above Frame Rate Frequencies

Dr. Steven M. Shepard* and Mr. David T. Sass
U.S. Army Tank-Automotive Command
ATTN: AMSTA-RSA
Warren, Michigan 48397-5000

INTRODUCTION

The use of infrared imaging for military applications is widespread and well documented. Although a great deal of progress has been made in improving the spatial quality of IR images, little attention has been paid to extracting the temporal information contained within an image. In fact, a single video frame from an IR imager is an RMS average of the temperature of the object field. For scenes containing time varying thermal phenomena, this type of averaging results in the loss of most temporal information. For example, a scene which contains spatially identical targets, where some of the targets have rapidly (>30 Hz) oscillating temperatures, will yield an IR image in which the oscillating and non-oscillating targets will be indistinguishable.

The desire to acquire more temporally precise IR image data prompted us to look more carefully at our imaging systems, in the hope that we could somehow improve the situation. In particular, we hoped to be able to adapt temporal signal processing operations commonly done on one dimensional signals, such as lock-in amplification and boxcar averaging, to two dimensional video signals. We agreed that an acceptable system should fall within the following constraints:

- 1) The system could not require the imager to be altered in any way.
- 2) The system output must be compatible with the standard (RS-170) video format, so that it could be interfaced with our existing hardware and software.
- 3) The system should be able to image repetitive events at above frame rate frequencies without aliasing.

SHEPARD AND SASS

- 4) The system must be portable, and suitable for use in a field test environment.
- 5) The system must be economical.

Analysis of our imagers led us to the following conclusions:

- 1) In order to enhance the temporal aspect of our image data, we would have to acquire phase information to accompany the intensity represented by each pixel. Thus we would need a reference signal temporally correlated to the target, analogous to the reference signal received by a lock-in amplifier, or the trigger signal used with a boxcar averager.
- 2) By selectively acquiring data according to its phase, we could build a more temporally precise image, in exchange for increased acquisition time.
- 3) The problem of aliasing at above frame rate frequencies could be solved by regarding our optomechanical scanning imager as a multiplexed line sampler, rather than as a frame sampler.

Working within these constraints, we developed a system which allows us to acquire thermal images with significantly higher temporal resolution than conventional imaging systems. Our technique is adaptable to a variety of imagers, both IR and visible, and is easy to implement, requiring only a microcomputer equipped with a frame grabber, and a small amount of dedicated hardware and software. In the following sections, we will describe the theory, operation, and application of the system, as well as some applications in which it has been most useful.

PROBLEMS IN THERMAL IMAGING

The two principal problems in conventional thermal imaging which were of concern to us are 1) lack of temporal resolution, and 2) aliasing. These problems are inherent to most imaging systems currently in use. The lack of temporal resolution is problematic in imaging both high and low frequency phenomena, while aliasing prevails only at frequencies greater than the video frame rate.

Lack of temporal resolution can be best described by analogy to a still photographic camera, which has an adjustable shutter speed. If one attempts to photograph a moving object with too low a shutter speed, the image of the target will appear blurred. To "freeze" a rapidly changing event, a fast shutter speed is used.

In the case of a video imager, the "shutter speed" is fixed, as it is determined by the video frame rate. We can define an "effective shutter

SHEPARD AND SASS

speed" (i.e. the integration time) as the time between the acquisition of the first and last pixels displayed in a given video field. For a standard (RS-170) imager, the integration time is 16.66 milliseconds. For an optomechanically scanned IR imager, the situation is usually more complex, so that the effective shutter speed is at least that long.

The presence of blurring in a visible image is obvious, as edges become indistinct. In an infrared image, the effects of low temporal resolution are harder to detect, as stationary targets with rapidly changing temperatures (or emmittances) will be thermally indistinct, although the geometric features may appear sharp and distinct. The overall effect is that by simply looking at the acquired IR image, there is no way of knowing whether or not information has been lost. IR images are further complicated by the fact that a single video field may be a combination of horizontal lines acquired from one of several detectors (for a scanning system using a line array detector), lines acquired by sweeping the object field in opposite directions (for a bidirectional scanning system), and lines which are duplicates of previously acquired lines (to account for the fact that the rate at which horizontal lines are scanned may be slower than the rate at which they are output to the video display system).

Our approach to improving the temporal resolution of our imager has been to selectively acquire data one horizontal line at a time. The onset of each horizontal line sweep of our imager is electronically compared to an external reference signal which is correlated to the repetitive target event. When the line onset and the reference signal have some user determined phase relationship, that line is stored in the system microcomputer memory. Multiple acquisitions of the same line result in the average of all samples of that line being stored. When a complete image has been acquired, i.e. the entire display has been filled, the RAM image is transferred to the display.

This same approach is used to overcome the aliasing which occurs when one attempts to image repetitive events which occur at above frame rate frequencies. Aliasing is the result of an insufficiently high sampling rate. By considering a video imager to be a frame (or field) sampling device, one is constrained to imaging targets at sub-frame rate frequencies. However, if data is acquired one line at a time, the imager is essentially being used as a line sampling device, and the sampling rate would correspond to the horizontal line sweep frequency of the imager. In the case of a conventional imager the line sweep frequency (and thus the sampling rate) is 15.75 kHz, while the frame rate is 30 Hz. In this case, using the "line-by-line" imaging mode would allow repetitive events at frequencies as high as 7.825 kHz (application of the Nyquist criterion requires that the sampling rate be at least twice as high as the event under consideration) to be imaged without aliasing, while conventional frame based imaging places an upper limit of 15 Hz on repetitive phenomena that can be imaged without aliasing.

SHEPARD AND SASS

HIGH SPEED IMAGING SYSTEM

Our system was built using an off-the-shelf Inframetrics IR-600 imager, and an IBM-AT compatible computer equipped with a GRAM frame grabber board by THERMOTEXNIX. No alterations were made to the imager. Instead, we developed a hardware interface and software driver to achieve the necessary synchronization for constructing images in the line by line mode. The hardware monitors the phase between the horizontal scanning mirror motion of the imager and a reference signal which corresponds to the timing of the thermal event of interest. When the user-prescribed phase relationship between these events is achieved, the appropriate line number is latched and transmitted to the host microcomputer after the entire field has been acquired. The software receives a signal indicating that a "good" event (i.e. desired phase relationship between horizontal scan and thermal event) has occurred, and freezes the field at the next vertical blanking interval. The good lines are copied from the frozen display to a RAM buffer where they are added to previously acquired lines. When a full set of image data (200 lines) has been collected (i.e. when all active lines have been sampled some user-specified number of times), a linear average (the sum of good lines for a particular line divided by the number of occurrences of good lines for that line) of the image is displayed. The net result is that signal-to-noise improvement can be achieved with no loss of temporal resolution.

The line-by-line approach to high speed imaging can be adapted to most video imagers, but it is particularly well suited to a bi-directionally scanned, optomechanical system such as the Inframetrics IR 600. In this imager, the horizontal scanning mirror is driven sinusoidally at 4 kHz, yet video line data is sent out at 15.75 kHz. To achieve compatibility with the RS-170 format, data is taken on the forward and rear passes of the mirror scan, making the effective scan rate 8 kHz. Lines scanned in the reverse direction are digitized, and reversed digitally before being re-converted to analog. To compensate for the difference in frequency between the scanning mirror and the video horizontal drive, each acquired line is sent out twice.

The IR 600 produces a high-quality image, which affords good spatial resolution, yet minimizes the lag and flicker common to many IR imagers. Unfortunately, any temporal interpretation of images from the IR 600 is complicated by several factors:

- 1) A given line may be a duplicate of the previous line.
- 2) A given line may be a digitally reversed version of a backwards-scanned line.
- 3) A given line may be both of the above.
- 4) A given scanned line may be displayed in either an odd or even video field.

SHEPARD AND SASS

The net effect of all of these factors is that the imager effectively has an 8:1 interlace, in the sense that it takes 8 video fields for a given line to cycle through the possible states determined by factors 1-4. Our system is able to acquire a temporally well-resolved image by synchronizing data acquisition to the actual motion of the horizontal scanning mirror. Operating in this manner, we are able to eliminate all duplicate lines, and to specify whether the image should be composed of scans in the forward, reverse, or both directions.

The Inframetrics IR 600 imager has an 8 kHz horizontal sampling frequency (4 kHz bi-directional horizontal scan rate), thus the effective shutter speed of the imager, when used in the line-by-line mode, is 125 microseconds, i.e. the upper limit for detection has been increased by a factor of 200 when compared to a conventionally obtained frame image.

APPLICATION

We have found the high speed IR imaging system to be a useful tool in a number of applications, including feature extraction/clutter reduction and vehicle signature analysis. We are currently using the system to study piston temperatures in a spark ignited engine at the University of Wisconsin Engine Research Center. In this particular application it is the enhanced temporal resolution of the system that is of critical importance. Although the engine speeds involved are not particularly high, the thermal phenomena under consideration are very fast with respect to engine rpm (e.g. the propagation of the flame across the cylinder, or the diffusion of heat on the piston head). Thus even if the engine speed is sufficiently low to allow a conventional image to be acquired without aliasing, the lack of temporal resolution in a conventional image will result in these faster thermal phenomena being "washed out".

The test engine at U-W was a single cylinder 4-stroke engine equipped with a ZnS viewing window to allow long (8-12 μm) radiation to pass, and a crankshaft encoder which allowed the exact angle of engine rotation to be monitored. In figure 1, a series of images shows the thermal state of the piston head during the first 21.2 degrees of crankshaft rotation (relative to top dead center). The top row shows the pre-ignition state of the engine from top dead center to 3.9 degrees (left to right). The 2nd through 4th rows show the propagation of the flame across the piston head (the flame appears as the brightest part of the image). The flame enters from the upper left region of the cylinder and moves diagonally across. Each image shown in the figure is a 125 μs "snapshot" of the engine acquired with the line by line technique.

CONCLUSION

The "line-by-line" imaging technique is an effective method for

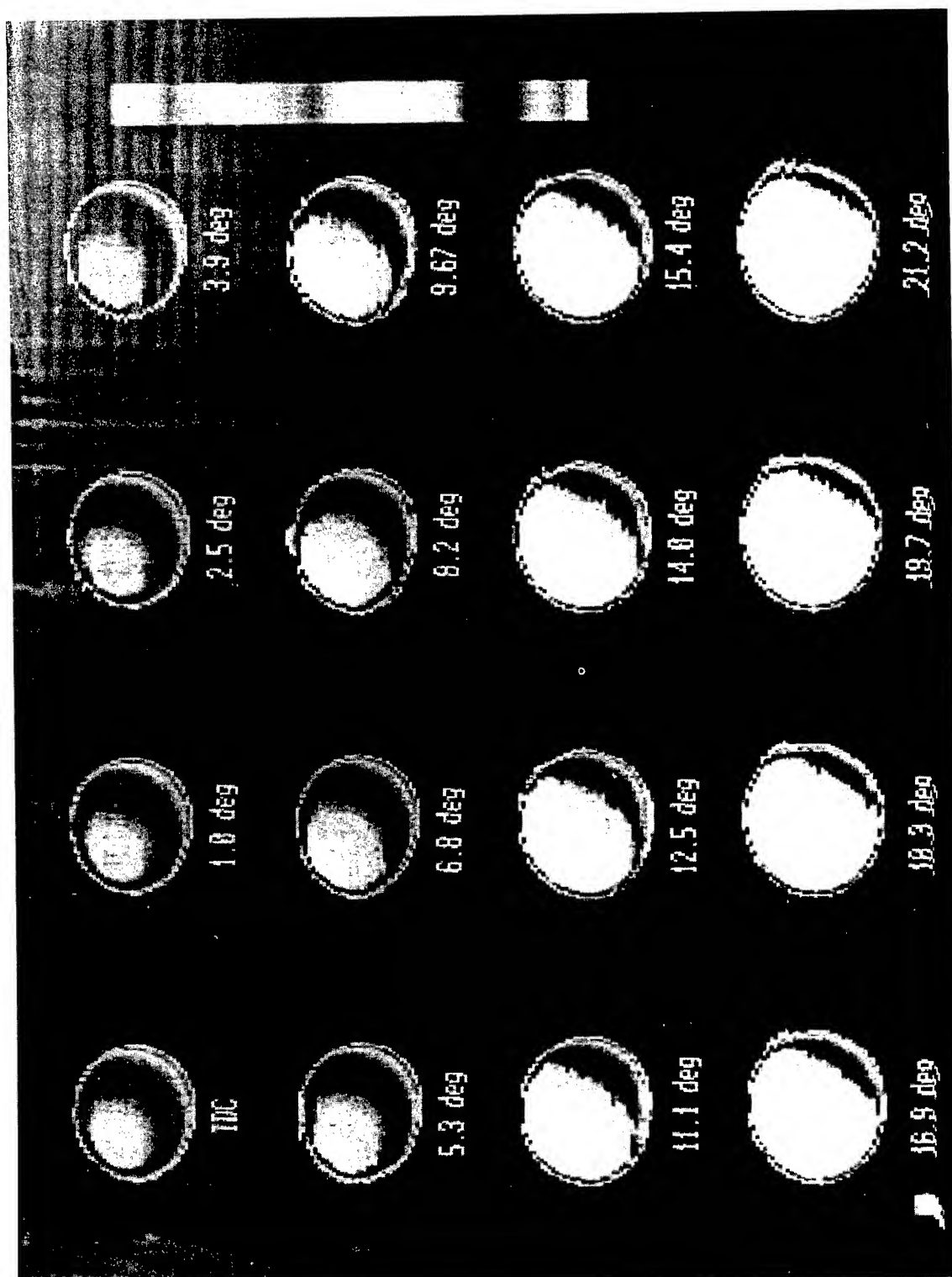


Figure 1. Temporally resolved thermal images of a piston.

SHEPARD AND SASS

observing repetitive phenomena at above frame rate phenomena, and for enhancing temporal resolution in both high and low frequency domains. The technique is particularly effective for optomechanically scanned, single element IR detectors, although it can be used with other imagers including visible and IR vidicons. Our concept of reducing the effective shutter speed, or integration time of the imager to achieve high speed capability and to enhance temporal resolution allows existing imagers to be used for these purposes without modification to the hardware.

Future work in this area will include the application of these techniques to focal plane array technology. We hope to be able to substantially decrease the effective integration time while also decreasing the acquisition times so that single shot events may be imaged with enhanced temporal resolution.

REFERENCES

1. S.M. Shepard and D.T. Sass, "Thermal Imaging of Repetitive Events at Above-Frame-Rate Frequencies", Optical Engineering, Vol. 29, No. 2, p. 105.
2. S.M. Shepard and D.T. Sass, "Feature Extraction in IR Images Via Synchronous Video Detection", in Proceedings of SPIE Thermosense XI, Vol. 1094, p.148, 1989.
3. S.M. Shepard, D.T. Sass, P. Kollmorgen, "An Enhanced Thermal Imaging Workstation", Proceedings of the Eleventh Annual Symposium on Ground Vehicle Signatures, August 1989, Houghton, Mich.

Optoelectronic Generation, Control, and Distribution of Microwaves (U)

*G. J. Simonis, K. Purchase, R. Hay, and N. Gupta
LABCOM, Harry Diamond Laboratories
2800 Powder Mill Road
Adelphi, MD 20783-1197

P. Ashley
US Army Missile Command
Redstone Arsenal, AL 35898-5248

1. INTRODUCTION

Phased array antennas perform critical functions in the Army Firefinder hostile-fire tracking radar and the Patriot air defense and missile interceptor systems. Such antennas also show promise for space-based communications, navigation, and radar applications. Conformal or "smart skin" antennas provide low protrusion and wind drag for aerial vehicle applications. However, size, weight, cost, and waveguide bandwidth and attenuation considerations are presently serious impediments to further implementation. Fiber optic distribution of microwave frequency, phase, and amplitude offers attractive features for these applications, especially where large antennas, large numbers of elements, and/or sophisticated antenna functions are involved.

The ongoing development of millimeter wave and microwave monolithic integrated circuits in the half-billion dollar/seven year DoD-MIMIC program will contribute substantially to the enhancement of the above antenna capabilities as well as many other Army microwave applications. However, the largest cost in the successful development of MIMIC technology will be the successful development of effective test procedures for the large number of MIMIC circuits to be produced. This is especially challenging at high microwave frequencies where microwave contact test probes are difficult to use. Optical means of injecting and sensing microwave signals offer a "contactless" method of testing MIMIC circuits. Other applications of optical signals used to introduce microwave modulation include the characterization of the frequency response of infrared detectors and the synchronization of high-frequency electronic clocks. Optical means can be used to send microwave signals covertly over large distances by means of fiber optic waveguides (less than 0.2 dB/km). Such transmissions are particularly challenging with conventional microwave techniques at high microwave frequencies. This capability may be brought to bear on applications such as anti-radiation missile counter-measures and remote antenna applications.

This paper reports the use of heterodyne techniques¹⁻³ using two ultrastable diode-laser-pumped Nd:YAG ring lasers⁴ to achieve frequency-diverse high-quality optical transmission of microwave signals, as well as III-V optoelectronics means of phase and amplitude control of these signals in integrated optics.⁵ We also report the use of an alternate approach involving Bragg cell and lithium niobate integrated optics technology for multichannel optoelectronic control of microwaves.

We believe that our results constitute the most attractive combination of discrete-frequency optical/microwave control characteristics reported to date, encompassing microwave quality, frequency tuning range, and range of amplitude and phase control. In the optoelectronic realm we believe that our demonstrated 42 dB amplitude modulation constitutes the largest modulation dynamic range ever reported in a III-V semiconductor waveguide. Also our optical heterodyne technique provides a relatively direct measure of optical phase change and index of refraction change in these waveguides when compared with most reported work. The lithium niobate optoelectronic approach demonstrates relatively mature technology which is approaching commercial availability.

2. OPTICAL HETERODYNE GENERATION OF MICROWAVES

The detection of optical signal modulation at microwave frequencies has been demonstrated up to equivalent frequencies as high as 100 GHz by means of fast optical pulses.⁶ Even higher modulation frequencies may be detectable by using one element as both an optical detector and mixer.⁷ However, the direct microwave modulation of a laser diode source or even the modulation of an optical signal with an external modulator becomes progressively more challenging as the modulation frequency enters this high-frequency microwave range (above 20 GHz).

Our approach to optically transmit a high-quality microwave signal is to superimpose two high-quality continuous-wave single-frequency optical signals which differ in frequency by the desired microwave modulation frequency. The mixing of these two signals on an optical square-law detector will provide the difference frequency as the output of the detector. Two lasers can have an extremely large difference frequency (easily tens of terahertz with some lasers). The microwave frequency-limiting factor is then the optical detector frequency response. The difference frequency is of extremely high quality if the optical sources are of high quality individually, or if they are coherently coupled to each other.^{1-3,8,9} This mixing approach to microwave modulation also lends itself to an attractive means of optoelectronic phase and amplitude manipulation, which is discussed in the next section.

For our optical mixing work, we use high-performance single-mode monolithic Nd:YAG crystal ring lasers (up to 3 mW each) optically pumped by diode lasers.^{1,4,10} The special features of the design of these lasers result in a highly coherent single-mode emission at 1.06 μm that is temperature tunable over as much as 100 GHz and is highly resistant to optical feedback and laser coupling. Similar performance is achievable with similar lasers at a 1.3- μm wavelength with different Nd:YAG absorption and lasing transitions. Fig. 1 shows the optical mixing configuration (as well as other components to be discussed below).^{1,11} The two optical beams are made collinear by means of beam splitters for each output beam combination desired. Indium gallium arsenide detectors having an active element diameter of 20 μm or less are used.¹² Selected detectors provide sensitivity well beyond their estimated 30-GHz 3-dB roll-off point because of the high signal level available. A Hewlett Packard 8566B spectrum analyzer is used for analysis.

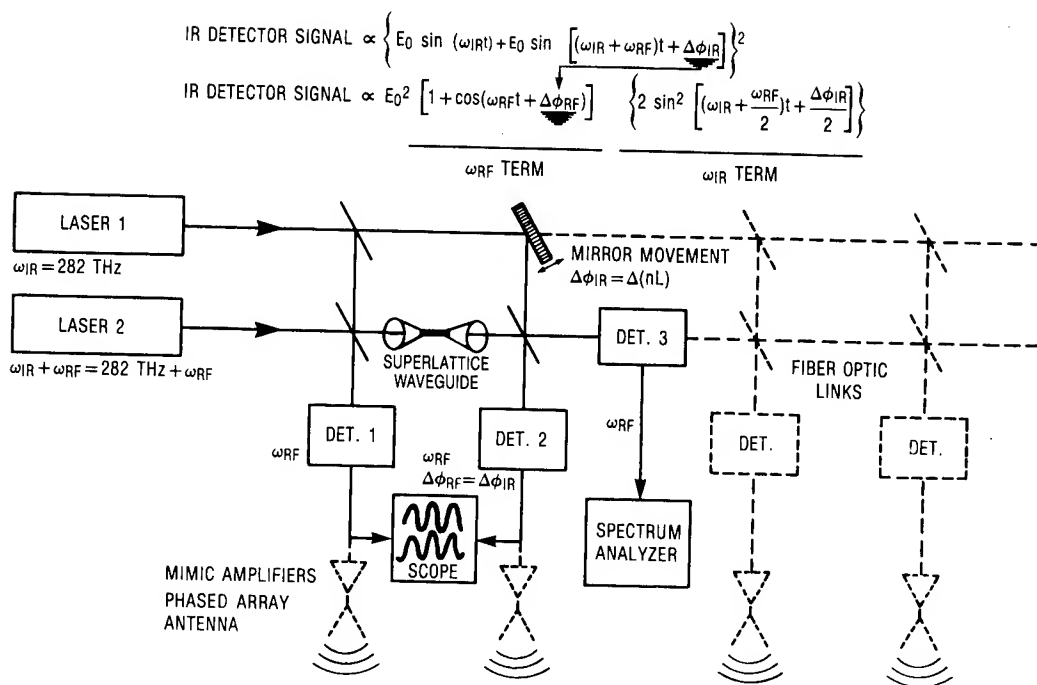


Fig. 1. Experimental heterodyne configuration. Induced optical phase/amplitude change results in microwave phase/amplitude change.

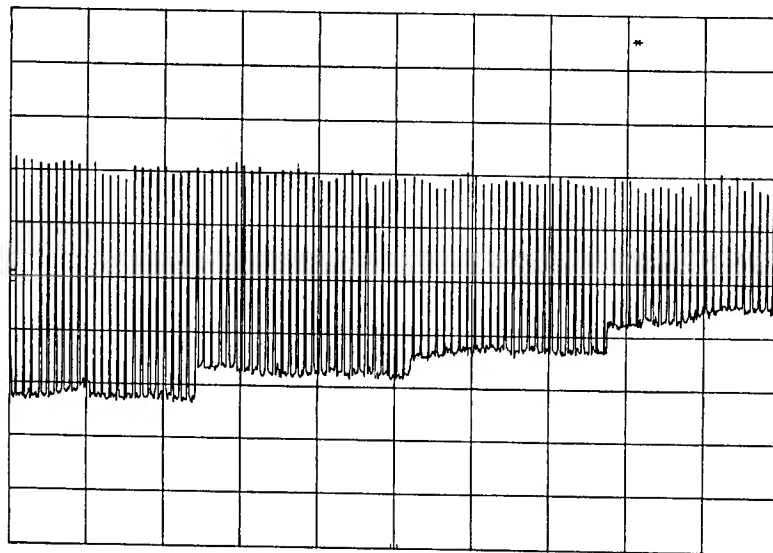


Fig. 2. Spectrum analyzer display showing heterodyne power as difference frequency is increased in steps, 2.4 GHz/div, 0 to 24 GHz, 10 dB/div, ref. level -10 dBm.

Fig. 2 presents a spectrum analyzer display of the amplified heterodyne signal as the difference frequency is advanced in steps and retained in the display. The small observed power changes with frequency are believed due primarily to microwave impedance mismatches and increasing microwave cable losses with frequency. The noise floor is spectrum analyzer limited in this low-resolution display. This heterodyne technique provides a very broad-band method of microwave sweeping that can be used to characterize infrared detectors and microwave components.

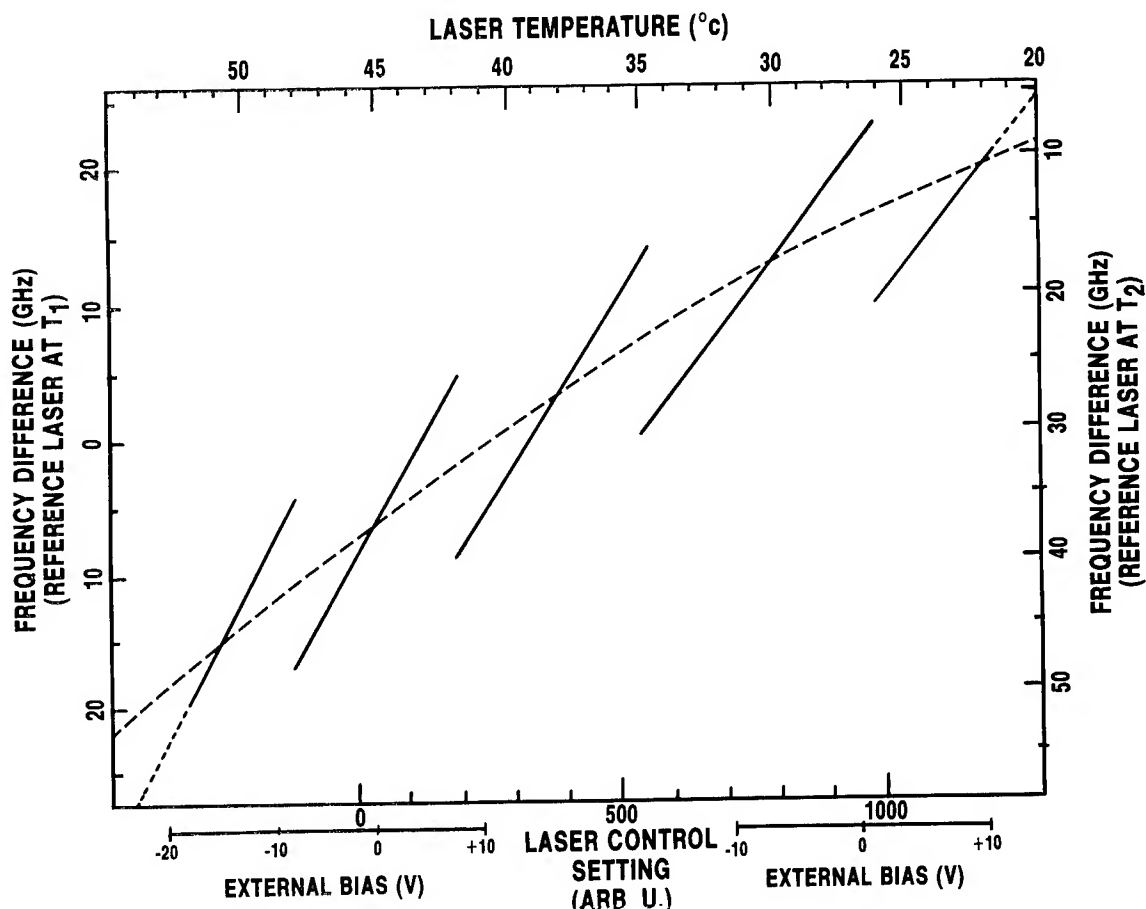


Fig. 3. Measured difference frequency between two lasers as one is temperature tuned relative to the second reference laser. Two vertical scales correspond to two different reference laser temperature (T_1 and T_2) and settings. Laser temperature control scale (arb. units) also incorporates external voltage override used to extend the temperature tuning range.

Fig. 3 shows the microwave difference-frequency selection that is achievable by temperature tuning the optical frequency of one of the lasers.¹¹ The tuning curve shows the difference frequency with the two vertical scales corresponding to two different frequency settings of the reference laser. A tuning range in excess of 50 GHz is demonstrated for a temperature tuning

range of 30°C. Frequency tuning of these lasers can be accomplished on a microsecond time scale over a smaller frequency range by inducing stress in the crystal with a piezoelectric element. An external mixer was added to the spectrum analyzer for the high-frequency measurements above 26 GHz. A doubling of the tuning range up to 100 GHz should be readily achievable by doubling the temperature tuning changes in the laser; such lasers are available. The dashed line in Fig. 3 shows the estimated frequency tuning of the Nd:YAG gain line center with temperature (1.0 GHz/°C), and the solid line presents the actual observed difference frequency tuning dictated by the temperature tuning of the cavity modes (3.1 GHz/°C). The horizontal temperature scale is approximate and somewhat nonlinear. Settings at which mode hops occur are apparent in the plot. The lasers have a tendency to operate on two cavity modes simultaneously when the laser is close to a mode hop, resulting in a self-beat signal near 14 GHz, but otherwise provide a single pure tunable heterodyne signal. Note that a 100-GHz tuning range for these 1.06- μm (282-THz) lasers would be only a 0.04-percent optical frequency tuning bandwidth.

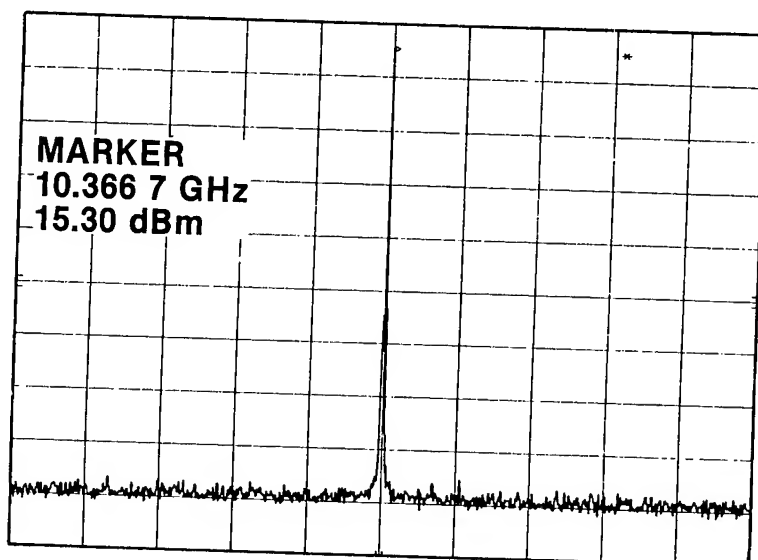


Fig. 4. Power spectrum of amplified heterodyne signal at 10 GHz, 10 MHz/div, 10 dB/div, ref. level 20 dBm, resolution bandwidth 3 kHz.

Fig. 4 shows the amplified heterodyne power spectrum at a difference frequency of 10 GHz and a scale of 10 MHz/div. The signal-to-noise ratio is 85 dB, and the signal appears highly coherent on this scale. Fig. 5 shows the same power spectrum with a frequency scale of 100 kHz per division, a resolution bandwidth of 3 kHz, and a sweep rate of 100 ms per division.¹¹ We find the 3-dB linewidth to be less than 5 kHz when the sweep time over the peak is of millisecond duration. The signal level at 300 kHz from line center is less than -115 dBc/Hz. Sidebands are clearly apparent about 50 dB below the main signal and about 200 kHz to either side of the main signal. These sidebands are due to resonant frequencies in laser amplitude fluctuations, referred to as relaxation oscillations in the individual Nd:YAG lasers. Laser

output leveling circuitry feeding back into the diode pump laser has been shown to suppress these sidebands by at least 20 dB.¹³ (Relaxation oscillations in diode lasers occur in the gigahertz range and are difficult to level out.) Our results involve no active laser frequency stabilization. Other researchers have recently reported locking the lasers to a Fabry-Perot cavity¹⁴ or to each other¹⁵ and have achieved 3 Hz and less than 1-mHz 3-dB heterodyne linewidths respectively.

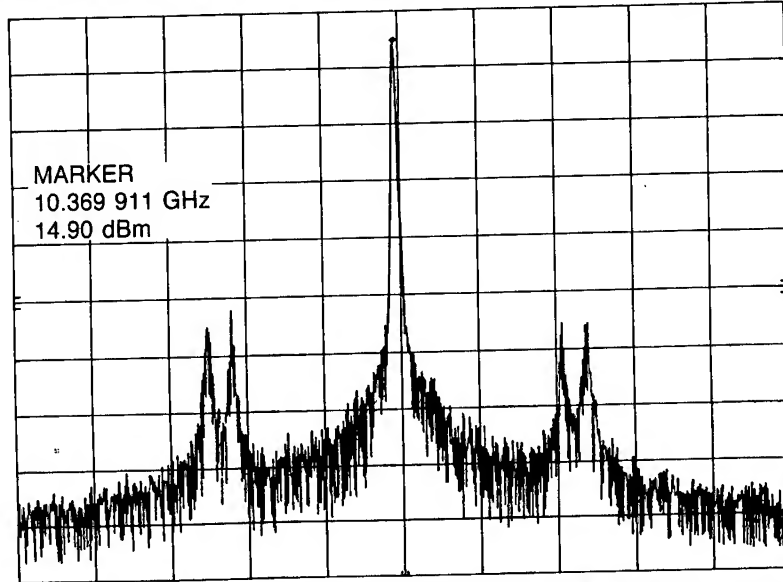


Fig. 5. Power spectrum of amplified heterodyne signal at 10 GHz, 100 kHz/div, 10 dB/div, ref. level 20 dBm, resolution bandwidth 2 kHz.

3. GaAs OPTOELECTRONIC CONTROL OF PHASE AND AMPLITUDE

In our present work, we are using a gallium arsenide doping superlattice in an optical waveguide to apply the optical amplitude modulation as portrayed in Fig. 6. The 1.06- μm below-band-gap laser signal to be controlled is focused into the 2- μm -thick GaAs superlattice core of the 1-mm-long waveguide with a 36 \times microscope objective. The GaAs core region is clad above and below by 1.5 μm of lower-index $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ to confine the propagating optical mode by total internal reflection. The superlattice has electrical contacts selectively accessing the interleaved p-doped and n-doped layers. (The superlattice and waveguide are discussed in more detail elsewhere.^{5,16}) The complex index of refraction can be changed either by applied voltage or by above-band-gap optical injection of electron-hole pairs. Optical injection basically counteracts any applied reverse bias and drives the superlattice into a forward bias condition as the intensity is increased.

The modulation of the amplitude of a heterodyne signal at 12 GHz as a function of voltage applied to the doping superlattice is shown in Fig. 7. A value of 2.7 V of reverse bias applied to the waveguide superlattice induces 42 dB of attenuation of the heterodyne signal. The observed irregular shape of the attenuation-versus-voltage curve is believed due to optical

interference between multiple waveguide modes and multiple reflections; the use of single-mode waveguides and anti-reflection coatings would minimize these effects. Fig. 8 shows typical behavior of the transmitted signal when milliwatts of above-band-gap chopped optical input is provided by an argon ion laser. The optically injected electron-hole pairs counteract the voltage-induced attenuation, resulting in an increased heterodyne signal.

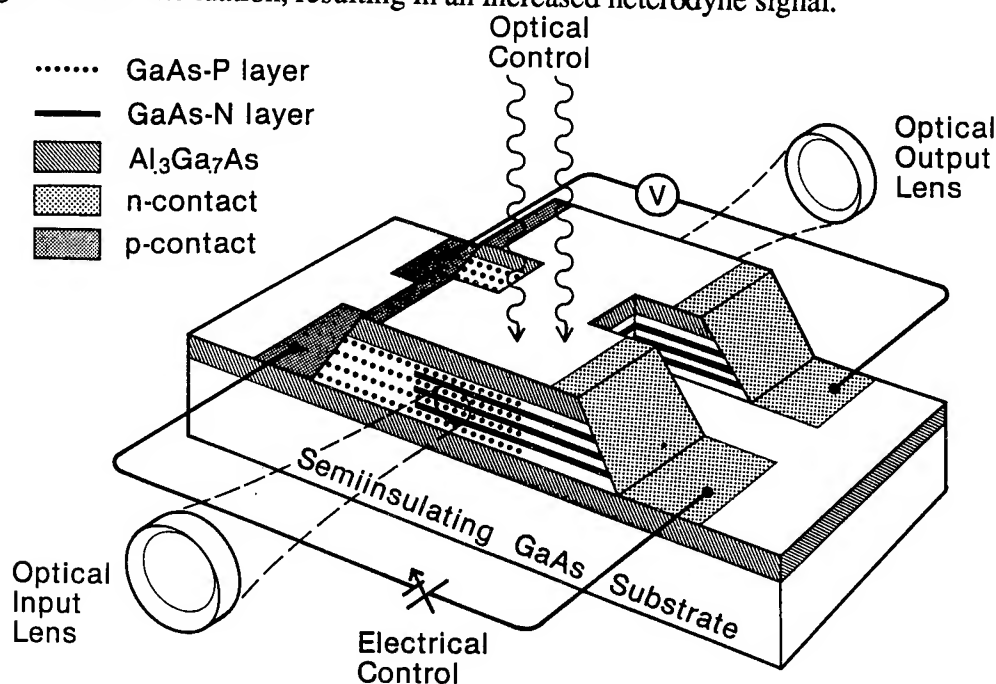


Fig. 6. Configuration of semiconductor optical waveguide, doping superlattice with selective contacts, and optical and electrical control signal inputs.

The phase shifting of a microwave modulation superimposed on an optical carrier is much more difficult to implement optically in a practical fashion than is the amplitude control of such a modulation, when thousands of parallel channels may be involved and the equivalent of five or more bits of phase shift control is needed for each channel. The optical path length (index of refraction times distance) changes that must be induced to phase shift the microwave modulation are on the order of the microwave wavelength in the fiber (millimeters to many centimeters or more). Most efforts to phase shift such microwave modulations involve switching optoelectronically between lengths of fiber optical waveguide. This involves many fibers and switches for each channel. This approach does have the merit of providing "time-delay" phase shifters; i.e., the phase shifter automatically provides wavelength-dependent phase shifts to correct for wavelength-dependent beam direction from an antenna array. This time-delay phase-shift feature is important for instantaneous broad-band antenna applications.

The alternative approach that we are using is to employ optical mixing to achieve the microwave modulation and to phase shift one of the optical carriers before the mixing is done. An optical phase shift applied to one of the optical carriers results in a corresponding microwave phase shift in the microwave modulation.^{1,2,8,11} Only an optical path length

change corresponding to the *optical* wavelength (micrometers of length change) needs to be introduced in order to change the relative phases of the microwave outputs by many π . The required changes in optical path lengths of micrometer dimensions are much more compatible with available electro-optic and integrated optical techniques. A demonstration of optical phase shifting employing acousto-optic Bragg cells for phase shifting control by optical beam steering has been reported.⁸ That configuration includes an optical means of "computing" the phase shift appropriate to each channel.

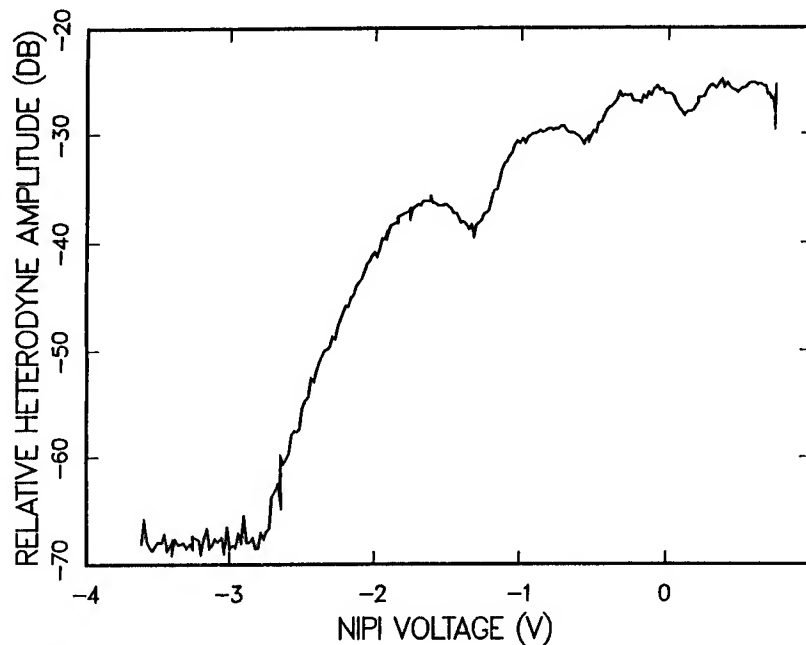


Fig. 7. Voltage-tuned power control of a 12-GHz heterodyne signal. Voltage is applied to a 1.2-mm-long doping superlattice (NIPI) waveguide.

Optical phase shifting may be achieved in our configuration either by mechanical mirror movement¹ or by optoelectronic means, as illustrated in Fig. 1. The optoelectronic semiconductor waveguide used for the phase shifting is the same device as used in the amplitude control and presented in Fig. 6. Fig. 9 shows an oscilloscope trace of two channels of heterodyne output with the scope trace triggered by the first channel. The use of a small (25-MHz) heterodyne difference frequency allows us to visualize directly the relative "microwave" phase of the second channel on the oscilloscope. Similar phase shift performance can be extended to higher frequencies. The phase of the lower heterodyne signal in Fig. 9 is controlled by the superlattice waveguide. The optical injection of carriers into the waveguide by milliwatts of argon ion laser was chopped, resulting in two different phases, apparent in the time-exposure picture. Using this oscilloscope technique we can measure electrically or optically induced phase shifts without the need for a microwave vector network analyzer. The extension of these results to higher frequencies could provide a high-frequency (up to 100 GHz or higher) vector network analyzer having fiber-optic interconnects. Fig. 10 presents the

electrically induced phase shift as a function of applied voltage. A voltage range of 3 V provides a phase shift control of 8π . The measurable voltage range was limited by increasing attenuation with increasing reverse bias. Several π of phase shift with very modest attenuation can be achieved by using forward bias and slight reverse bias. The slope of the voltage-controlled phase is relatively constant when compared with the slope of the amplitude control of Fig. 7.



Fig. 8. Modulated transmission of reverse biased waveguide as it is periodically optically injected with above-band-gap photons to increase below-band-gap transmission, 10 ms/div.

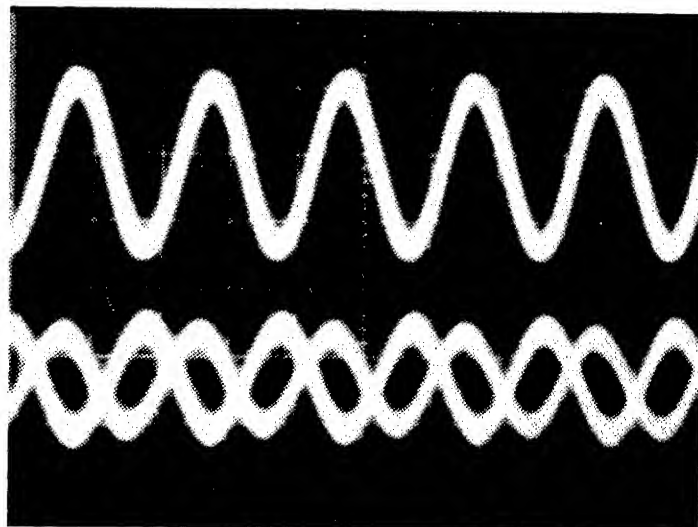


Fig. 9. Two heterodyne channels at 25-MHz difference frequency. Lower channel phase shifted/unshifted by optical injection of carriers into superlattice waveguide 20 ns/div.

4. LITHIUM NIOBATE OPTOELECTRONICS IMPLEMENTATION

We are also working with lithium niobate optoelectronics for laser combining and control. Such optoelectronics is more mature than GaAs optoelectronics and is currently becoming commercially available, including single-mode fiber optic coupling. High-efficiency optical phase shifting can be readily achieved in lithium niobate.¹⁷ Our experimental configuration is presented in Fig. 11. Our initial efforts employed a helium-neon laser at a $0.63\text{-}\mu\text{m}$ wavelength.¹⁸ We are presently using a single-transverse-mode temperature-stabilized semiconductor diode laser at a $0.83\text{-}\mu\text{m}$ wavelength.¹⁹ Our waveguide power splitter/phase shifter was fabricated by thermal in-diffusion of titanium into x-cut lithium niobate. The waveguides are $3.5\text{ }\mu\text{m} \times 2.5\text{ }\mu\text{m}$, are Y-propagating, and are spaced $50\text{ }\mu\text{m}$ apart at the output. They are single mode waveguides supporting either TE or TM propagation. There is an optical phase shifter on each of the eight output waveguides.

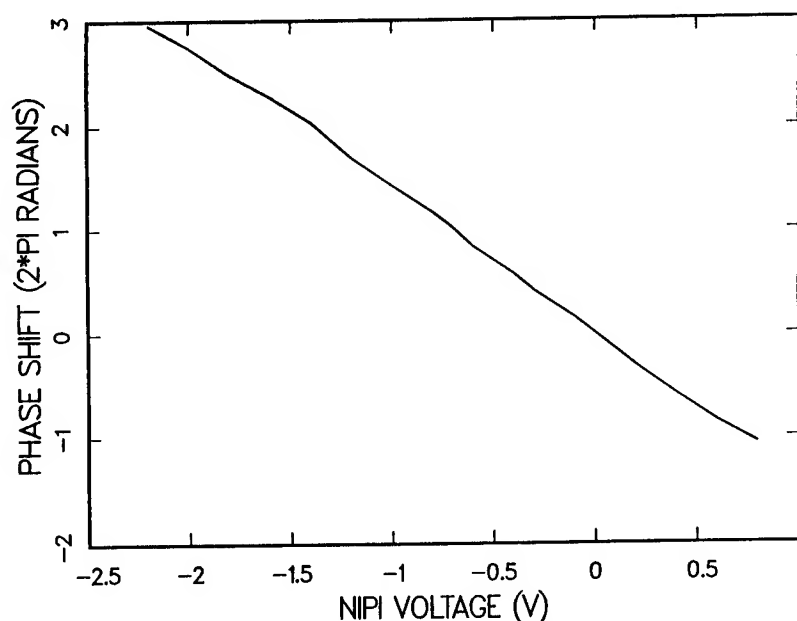


Fig. 10. Phase change induced in one heterodyne channel relative to another by voltage applied to the optoelectronic superlattice waveguide.

As an experimental expedient, we used a cylindrical lens to combine a frequency-shifted reference signal with the waveguide outputs since we did not have a second functional waveguide power splitter available. This is a very inefficient power combiner, but we nonetheless had sufficient optical power available in the frequency-shifted arm of the configuration to saturate the detectors. The optical difference frequency of 100 MHz is obtained by frequency shifting one of the beams with a TeO_2 Bragg cell optimized for this wavelength range. Further details of the experimental configuration are presented elsewhere.¹⁹

Fig. 12 shows a CCD camera image of the waveguide outputs with and without the frequency-shifted reference signal superimposed. Good overlap is provided between the frequency-shifted signal and two of the waveguide outputs. Fig. 13 shows heterodyne spectra for two of the channels obtained with an HP8590A spectrum analyzer. The signal-to-noise ratio was insufficient to allow direct phase visualization, as compared to the process of combining these two channels into a Mach-Zehnder-like interferometer to compare phases between channels. Stray signals, severe insertion losses, and limited signal-to-noise ratio are causing experimental problems, in our present setup. However, direct efficient coherent fiber optic coupling within and between signal channels will make dramatic improvements in overall performance. It is clear that currently available lithium niobate/fiber optic technology can provide a means of generating and controlling multiple microwave channels, at least up to the 3 GHz state-of-the-art of Bragg cells. The use of harmonics or the two-laser heterodyne approach can readily extend the frequency range of the lithium niobate approach.

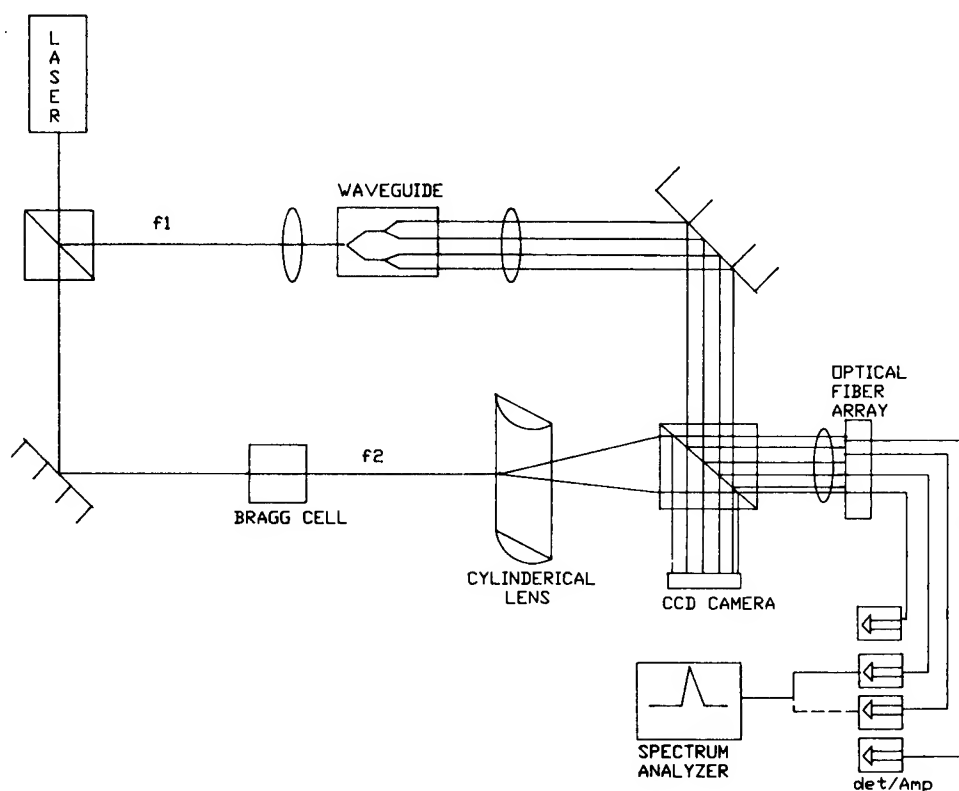


Fig. 11 Experimental configuration using lithium niobate integrated optical waveguide power splitter with phase shifter on each channel.

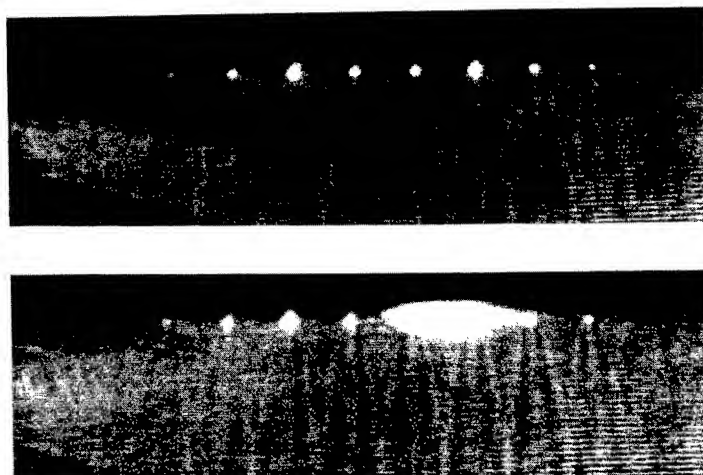


Fig. 12 CCD camera image of waveguide signals a) waveguide output only b) waveguide output plus frequency-shifted signal from Bragg cell.

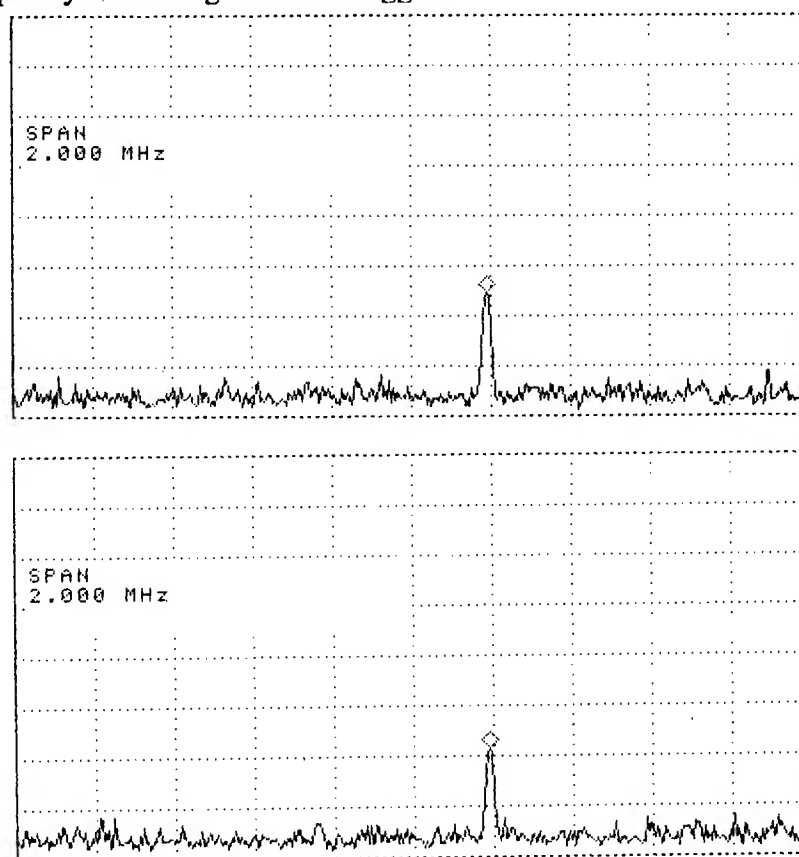


Fig. 13 Heterodyne spectra at 100 MHz for two channels in configuration shown in Fig. 11, 10 dB/div., reference level 0 dBm, 200 kHz/div, resolution bandwidth 10 kHz.

5. CONCLUSION

We have demonstrated the use of optoelectronic heterodyne techniques to generate high-quality microwave signals from dc to 52 GHz which can be distributed over fiber-optic lines. We also have presented the use of III-V optoelectronic waveguides to induce phase changes of more than of 8π and amplitude changes of more than 40 dB. Both electrical and optical input control mechanisms have been used. We have also demonstrated the use of lithium niobate optoelectronics and Bragg cell technology for the control of multiple microwave channels. These approaches are compatible with high-density optoelectronic integration, which is required for the large number of channels in a phased-array antenna. These techniques are also attractive for other optical feed applications, such as a vector network analyzer, and for the characterization of the frequency response of infrared detectors.

6. ACKNOWLEDGEMENT

Special thanks goes to Gulam Hasnain of AT&T Bell labs, Murray Hill, NJ, who provided the selectively contacted doping superlattice samples from his Ph.D work at the University of California, Berkley, CA.

7. REFERENCES

1. G. J. Simonis and K. G. Purchase, "Optical generation and distribution of microwave signals," Thirteenth Intl. Conf. on IR and Millimeter Waves, R. J. Temkin, editor, Proc. SPIE, Vol. **1039**, pp. 102-103, 1988.
2. R. A. Soref, "Voltage-controlled optical/RF phase shifter," J. Lightwave Tech., Vol. LT-3, pp. 992-998, 1985.
3. D. K. Donald, D. M. Bloom and F. K. David, "Efficient, simple optical heterodyne receiver: DC to 80 GHz," Optical technology for Microwave Applications II, SPIE, Vol. **545**, pp. 29-34, 1985. H. Fetterman, C. Liew, and W. L. Hgai, "Millimeter-visible injection locking and testing," *ibid*, pp. 26-28.
4. T. J. Kane and R. L. Byer, "Monolithic unidirectional signal-mode Nd:YAG ring laser," Opt. Lett., Vol. **10**, pp. 65-67, 1985.
5. G. J. Simonis and K. G. Purchase, "Optical control of optical phase and amplitude in a doping superlattice waveguide," Conference on Lasers and Electro-optics, 1989 Technical Digest Series, Vol. **11**, Optical Society of America Washington, DC, 1989, Th J3.
6. S. Y. Wang and D. M. Bloom, "100 GHz bandwidth planar GaAs Schottky photodiode," Electron Lett., Vol. **19**, pp. 554-555, 1983.

7. C. Rauscher, L. Goldberg, and A. M. Yurek, "GaAs FET demodulator for optical-microwave links," *Electron. Lett.*, Vol. **22**, pp. 705-706, 1986.
8. M. Tamburrini, M. Parent, L. Goldberg, and D. Stillwell, "Optical feed for a phased array microwave antenna," *Electron. Lett.*, Vol. **23**, pp. 680-681, 1987.
9. L. Goldberg, A. M. Yurek, H. F. Taylor, and J. F. Weller, "35GHz microwave signal generation with an injection-locked laser diode," *Electron. Lett.*, Vol. **21**, pp. 814-815, 1985.
10. A. C. Nilsson, E. K. Gustafson, and R. L. Byer, "Eigenpolarization theory of monolithic nonplanar ring oscillators," *IEEE J. Quantum Electron.*, Vol. **QE-25**, pp. 767-790, 1989.
11. G. J. Simonis and K. G. Purchase, "Optical generation, distribution, and control of microwaves using laser heterodyne," *IEEE Microwave Theory Tech. MTT-38*, 1990, (to be published, May 1990). G. J. Simonis, R. G. Hay, and K. G. purchase, "Optical control of microwaves with III-V semiconductor optical waveguides," *Optoelectronic signal processing for phased-array antennas II*, SPIE Vol. 1217, 1990.
12. J. Schlafer, C. B. Su, W. Powaznik, and R. B. Lauer, "20 GHz bandwidth in GaAs photodetector for long-wavelength microwave optical links," *Electron. Lett.*, Vol. **21**, pp. 469-471, 1985.
13. Lightwave Electronics Corp., private communication, 1988.
14. T. Day, E. K. Gustafson, and R. L. Byer, "Active frequency stabilization of a 1.062 μm , Nd:GGG, diode laser pumped nonplanar ring oscillator to less than 3 Hz of relative linewidth" (submitted to *Optics Letters*).
15. K. J. Williams, R. D. Esman, L. Goldberg, M. Dagenais, and J. F. Weller, "Active offset phase locking of Nd:YAG 1319-nm nonplanar ring lasers to 34 GHz," *Electronic Lett.*, Vol. **25**, p. 1242, 1989..
16. G. Hasnain, C. Chang-Hasnain, A. Dienes, and J. R. Whinnery, "Characteristics of a monolithically integrated doping superlattice optical circuit," *Appl. Phys. Lett.*, Vol. **52**, pp. 1765-1767, 1988.
17. A. C. Alferness, L. L. Buhl, M. D. Divino, S. K. Korotky, and L. W. Stutz, "Low-loss, broadband Ti:LiNbO₃ waveguide phase modulators for coherent systems," *Electronics Lett.*, Vol. **22**, pp. 309-310, 1986.

18. N. Gupta, G. J. Simonis, and P. Ashley, "Multichannel optical heterodyning using integrated optics," 1989 Annual Meeting of the OSA, Orlando, FL, Oct 1989.
19. N. Gupta, G. J. Simonis, and P. Ashley, "Lithium niobate integrate-optics demonstration of optical control of microwaves," Optoelectronic signal processing for phased-array antennas II, SPIE Vol. 1217, 1990.

**An Integrated Target Acquisition System
for Fire Control
and Autonomous Acquisition (U)**

Dr. S. Richard F. Sims*

U.S. Army Missile Command
Research Development and Engineering Center
AMSMI-RD-AS-SS
Redstone Arsenal, Alabama 35898-5253

ABSTRACT

Several target pattern recognition and cueing implementations have been mechanized previously based on classical statistical pattern recognition methodology. The performance of these techniques typically demonstrate a significant false alarm rate. The consumption by human operators or an autonomous missile system of significant false alarms has demonstrated a clear detrimental performance effect.

The development of synthetic discriminant functions (SDF) for the purpose of target acquisition, optimum aimpoint selection and non-cooperative target recognition on tactical missile applications emphasize the potential usefulness of these in an integrated approach. Current moving target indication (MTI) techniques have also shown significant potential in an integrated system for target detection when sensor motion is also present.

Design and development of an integrated automatic target recognition (IATR) system using both classical statistical methods and knowledge based methods synergistically to perform effectively under battlefield conditions have the potential of providing a significantly superior capability. Synthetic

discriminant functions with a model data base is considered plausible for the knowledge base. Parallel design implementation of these processes shall provide necessary synergism and high throughput in the design. A description of the synthetic discriminant function techniques and use of MTI with a composite IATR architecture is described.

A Composite Architecture for an Integrated Automatic Target Recognition System

Figure 1 illustrates an integrated approach to pattern recognition which allows performance improvement, under many conditions, over classical techniques. The application of this architecture has been toward strictly passive sensors, however, it is clear it is also appropriate for a sensor fused approach as well. The "center line" of Figure 1, which is the classical processing methodology, has been studied diligently for more than 15 years, and when implemented independently produces significant false alarms and misclassifications.

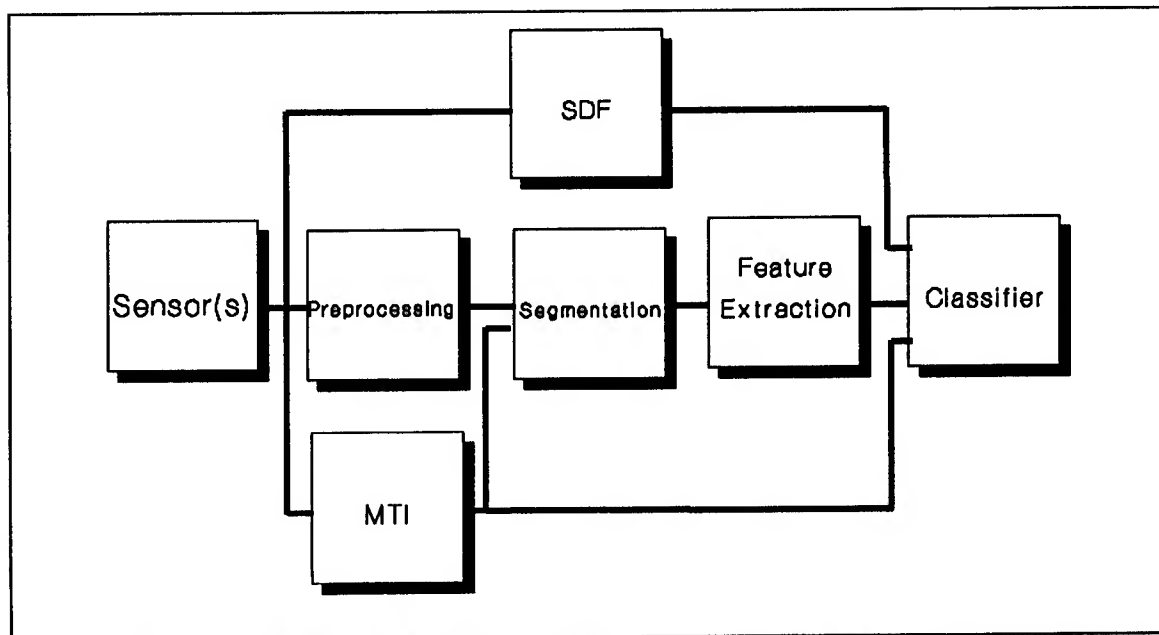


Figure 1 - Integrated Target Acquisition System Architecture

Recent methods for improving feature selection and classification such as genetic algorithms¹, reasoning with uncertainty, and neural nets have optimized the performance and training efficiency in these areas, however other synergism is required to produce significantly useful performance.

When targets are moving MTI produces relatively ideal segmentation and hence improves the classification, and just as important when no target motion is present MTI clearly does not produce anything to inhibit the classical methodology.

Synthetic Discriminant Functions (SDF) can be loosely described as multiplexed matched filters that are target range and aspect invariant. Every technique has limitations induced by dynamic range or resolution issues, however, when the SDF produces an output it provides a performance contribution that may not be obtainable from other means. SDFs are useful both in the detection and recognition of targets and therefore have a significant decision making capability.

Scenario Dependencies

Scenarios requiring automatic target acquisition vary significantly as to the amount of time allowed for processing the sensor information, and options to reprocess the data. For example, when time lines allow, the processed results generated from the architecture in Figure 1 on a full field of view sensor can then be used to establish a sequence of events which define subsequent processing to refine results and eliminate false alarms by controlling a smaller field of view onto selected potential target areas. Reprocessing a scene area from results generated in a large field of view subsequently by using a smaller field of view will provide more target information hence better performance than could otherwise be produced. Performance is therefore not strictly algorithm result dependant but in addition relates directly to the time lines associated with the processing throughput and sensor control capability.

Synthetic Discriminant Functions (SDF)

Synthetic Discriminant Functions have had several different names over the last ten years of their development from the earliest type called linear combinatorial filters to one of the latest versions called the minimum average correlation energy filter. The ten years of development produced many different variations and efforts toward significant advancement over two dimensional matched filters to allow scale and target aspect invariance.

Standard SDF using N training images d elements long where $d \gg N$.

$$\bar{h} = \sum_{i=1}^N a_i \bar{x}_i = \text{SDF filter}$$

$$\bar{h}^T \bar{x}_j = v_j$$

$$\sum_{i=1}^N a_i \bar{x}_i^T \bar{x}_j = v_j = \sum_{i=1}^N a_i R_{ij}, j=1,2,3, \dots, N$$

$$\bar{R} \bar{a} = \bar{v}$$

$$\bar{a} = \bar{R}^{-1} \bar{v}$$

The values of $v_j = 1$. The x_i are the training images.

Figure 2

SDFs can be categorized into three major stages of development each of which is considered an improvement over the previous one. The original or first SDF method is a linearly combined reference set² using the technique outlined in Figure 2³.

The severe limitation of this original approach is that the correlation surface is not guaranteed to be anything specific or defined except at the registered position. In other words there is no control over the output correlation surface except at one point. This result is of course in general much different than a matched filter would produce and was not very useful until phase encoding schemes were suggested.⁴ The phase only encoding, shown in Figure 3, and the binary phase encoding produce dramatically improved results.

The second category of SDF techniques is called the Correlation SDF⁵ and has an intuitive aspect which is easily visualized.⁶ The information used in building the SDF are the references themselves and noting the fact that the system of equations in Figure 2 is under determined the insertion of constants at desired places using a duplication of offset

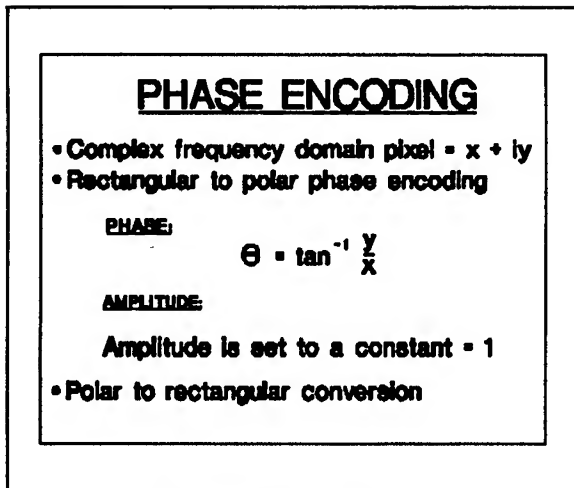


Figure 3

references creates specific control points on the correlation surface. This technique is used to create a distinctive correlation surface at and around the registered position. In addition this second method retains the limitation of the original SDF method in that it does not control the correlation anywhere except at the defined control points and the surface can take on uncontrolled values at all other positions. This uncontrollability is not necessarily a problem if a subsequent correlation stage uses an ideal result cross correlation to extract the unique surface shape sculpted by the control points selected. This approach does however add that additional computation. For other than single pass cross correlation for near term digital implementation this methodology performs better than the phase encoded methods and shows more potential detectability.

The third category of SDF development looks at a more comprehensive way of developing the SDF filters in that they strive to minimize the variance⁷ and then in another more easily computed version to minimize the average correlation energy (MACE)⁸. These methods vary significantly from the earlier techniques in that these approaches try to control the overall correlation surface by implicitly minimizing a global feature. An outline of the development of the MACE filter is shown in Figure 4 and Figure 5.

The MACE SDF produces better results than all other techniques described⁹, however, many researchers are working in this area and have developed hypotheses that may prove more robust.

Implementation of the SDF cross correlation is easily done in real time using state of the art digital VLSI electronics

$$g_i(n) = x_i(n) \odot h(n)$$

$$E_i = \sum_{n=1}^d |g_i(n)|^2 = (1/d) \sum_{k=1}^d |G_i(k)|^2 \\ = (1/d) \sum_{k=1}^d |H(k)|^2 |X_i(k)|^2$$

In vector form can also be written as

$$E_i = H^* D_i H$$

$$\text{where } D_i(k,k) = |X_i(k)|^2$$

and the correlation peak amplitude constraint is

$$g_i(0) = X_i^* H = u_i$$

Figure 4

$$E_N = (1/N) \sum_{i=1}^N E_i = (1/N) \sum_{i=1}^N H^* D_i H$$

$$\text{letting } D = \sum_{i=1}^N \alpha_i D_i \quad \text{if all } \alpha_i = 1$$

then

$$E_N = (1/N) H^* D H$$

By minimizing E_N subject to $X^* H = u$ using the method of Lagrange multipliers the solution for the MACE filter is

$$H = D^{-1} X (X^* D^{-1} X)^{-1} u$$

Figure 5

and has potential for direct optical implementation at higher rates.

Moving Target Indication (MTI)

Detection of target motion from a fixed sensor in real time is an achieved technological capability, however, when the sensor starts moving, scene dynamics for this process become very complex. Dynamic scene analysis provides an approach to the removal of sensor motion and consequently the basis for discerning relative target motion.¹⁰ The implementation of the image "flow field" extraction in real time allows moving target segmentation, as shown in Figure 7 and Figure 8. The architecture of Figure 1 indicates the MTI output to be inserted into the segmentation stage when available which in general will improve detection and classification. The false alarms produced strictly by motion detection

processes are generated on data sets that include windy days, other non-target moving objects, and fixed objects which have a different relative velocity field. These limitations can be

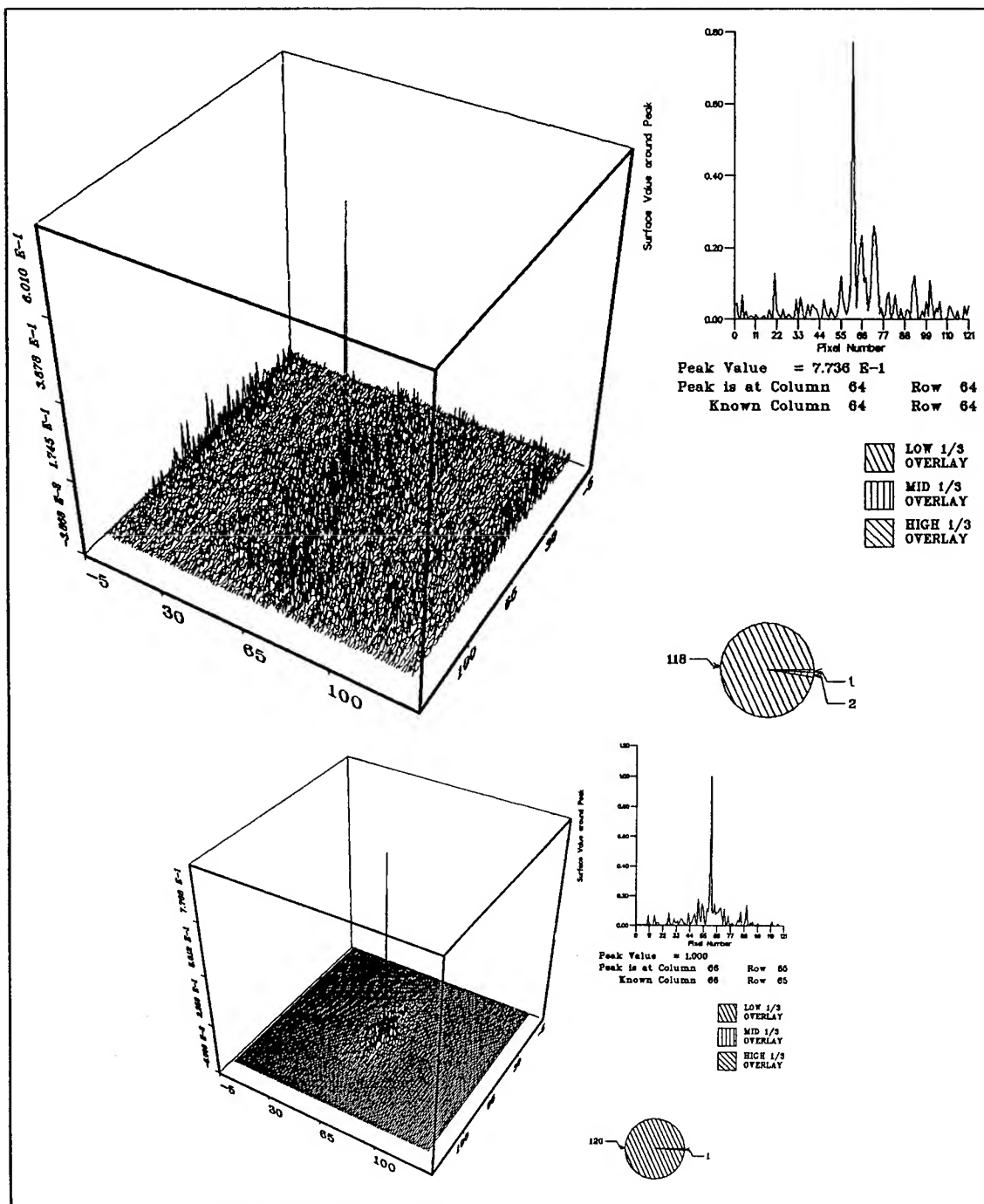


Figure 6 - SDF Correlation Surfaces



Figure 7 - Moving sensor with moving target

accommodated, as always, by operator adjustment of sensitivity parameters with knowledge of the effect on performance.

Autonomous Acquisition Issues

The application of the architecture of Figure 1 to an autonomous smart weapon poses questions related to the allowed levels of false alarms and the physical implementation limitations. In general it is clear that few, if any, false alarms can be tolerated for a smart weapon. In some cases the weapon must be able to perform the acquisition function and

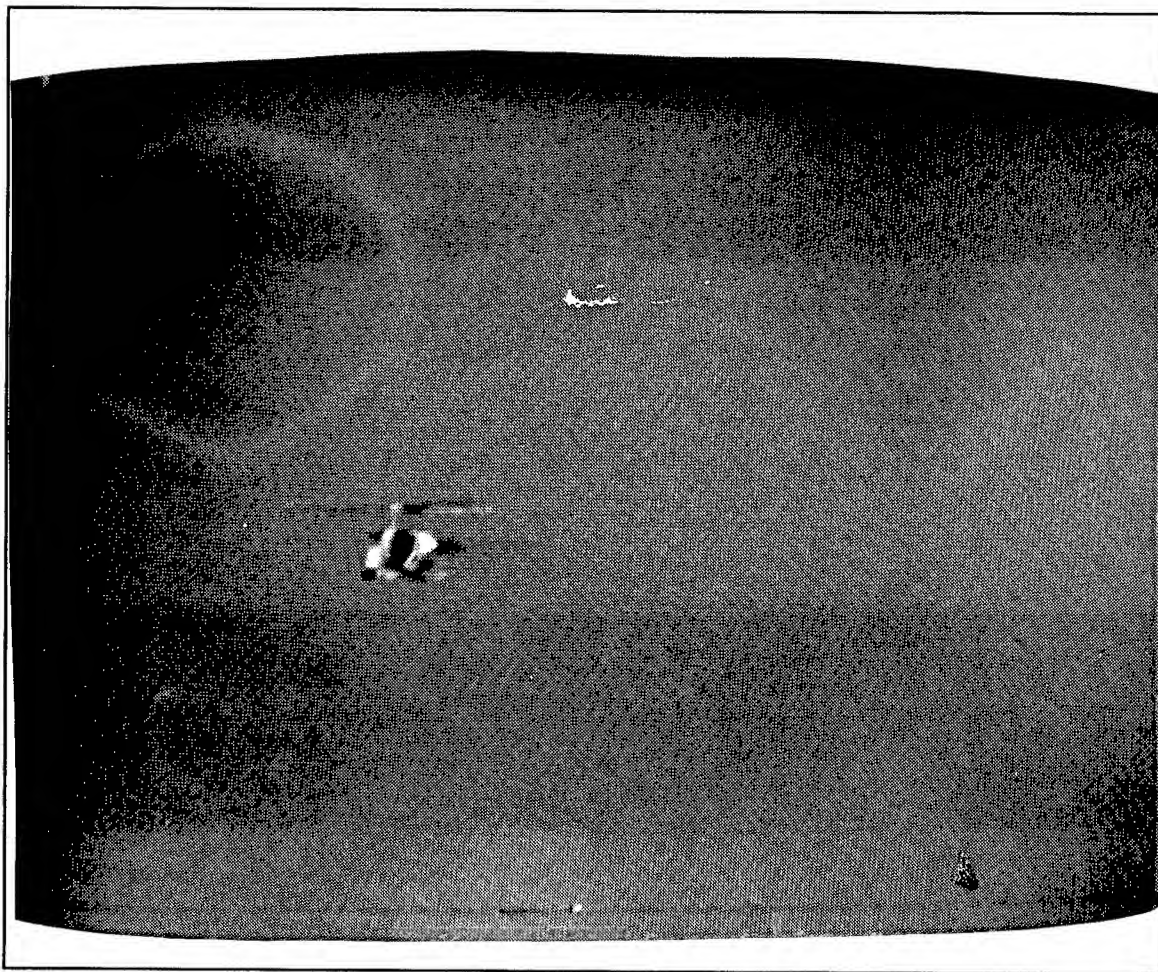


Figure 8 - Moving sensor with moving target and background flow field removed

then provide a reacquisition capability when obscuration temporarily occurs. In addition optimum aimpoint selection is required to provide maximum effectiveness. All of these functions are a requirement of and can be provided by the architecture in Figure 1. The adjustment of detection parameters of the classical methodology independently effects the probability of detection and false alarm rates, however in conjunction with the MTI and SDF processes can produce performance acceptable for restricted autonomous scenarios, where classical methods alone can not.

Implementation and the State of the Art

Implementation of the integrated ATR architecture into a generically usable configuration for smart weapons exceeds the state of the art currently, however, with the significant advance in optical and VLSI architectures now underway the potential for this is imminent. Fire control applications have less stringent needs and hence can accommodate the current technology for immediate implementation.

Flight hardware is currently being fabricated for the SDF function in Figure 1. These prototype boards will fit in the warhead void for the HELLFIRE missile during a technical flight demonstration. Real time implementation of the MTI function in Figure 1 is also currently being fabricated for testbed evaluation.

Hybrid implementation with optical and digital computing components will provide a substantial throughput and performance gain when obtainable. Optical processing of multiple parallel SDFs and digital processing of the classical techniques and MTI seems most suitable, at this time, as an eventual implementation.

Summary

When a sensor or sensors generate highly discriminative target information directly the amount of processing of that information for detection, recognition, and identification is clearly minimized, however, in a heavily countermeasured battlefield environment an integration of the most robust processes and sensors is necessary to provide superior performance.

1. **W. Siedlecki, J. Sklansky**, "A note on genetic algorithms for large-scale feature selection," Pattern Recognition Letters 10 335-347 (1989).
2. **C.F. Hester and D. Casasent**, "Multivariant Technique for Multiclass Pattern Recognition," Appl. Opt. 19, 1758 (1980).
3. **B.V.K. Vijaya Kumar**, "Optimality of Projection Synthetic Discriminant Functions," SPIE Vol. 579, Sept. (1985).
4. **Joseph L. Horner and Peter D. Gianino**, "Applying the phase only filter concept to the Synthetic Discriminant Function correlation filter," SPIE Vol 519 Analog Optical Processing and Computing (1984).
5. **D. Casasent and Wen-Thong Chang**, "Correlation synthetic discriminant functions," Appl. Opt. 25, 2343 (1986).
6. **E. Jenkins, J. Morris, S.R.F. Sims**, "Synthetic discriminant functions for target correlation," GOMAC Orlando, Florida (1987).
7. **B.V.K.V. Kumar**, "Minimum-variance synthetic discriminant functions," J. Opt. Soc. Am., Vol.3 No.10, Oct. (1986).
8. **A. Mahalanobis, V. Kumar, D. Casasent**, "Minimum average correlation energy filters," Appl. Opt. 26, 3633 (1987).
9. **S.R.F. Sims, J.A. Mills**, "Synthetic Discriminant Function Filter Performance Evaluations," SPIE Hybrid Image and Signal Processing II April (1990).
10. **P.J. Burt et al.**, "Object Tracking with a Moving Camera," IEEE Workshop on Motion, March (1989).

POLY(ORGANOPHOSPHAZENES) FOR NONLINEAR OPTICAL APPLICATIONS (U)

R. E. Singler, Dr., A. J. Jaglowski, Dr., M. S. Sennett, Dr.
Polymer Research Branch SLCMT-EMP
U.S. Army Materials Technology Laboratory
Watertown, MA 02172-0001

Reginald A. Willingham, Dr.
Physical Sciences Division STRNC-YMS
U. S. Army Natick RD&E Center
Natick, MA 01760-5020

INTRODUCTION

The use of polymers as materials for nonlinear optical (NLO) applications is gaining increasing attention because of the ability to tailor molecular structures which have inherently fast response times and provide second and third order molecular susceptibilities. Polymers provide synthetic and processing options that are not available with other classes of NLO materials, as well as excellent mechanical properties, environmental resistance, and high laser damage thresholds (1).

In the design and synthesis of NLO materials, liquid crystalline polymers offer some advantages over other polymer types. Side-chain liquid crystalline polymers are of particular interest due to the ease of synthesis and processing of these materials relative to main-chain liquid crystalline polymers. The concept of a side-chain liquid crystal polymer has been demonstrated in a number of laboratories and is well documented in the literature (2). Most of the side-chain liquid crystalline polymers reported to date contain polysiloxane, polyacrylate or polymethacrylate main chains. More recent studies on the effect of backbone flexibility now include the use of flexible poly(ethylene oxide) or more rigid poly(α -chloroacrylate) chains.

Phosphazene chemistry offers a new approach for design and synthesis of side-chain liquid crystal polymers and for the development of NLO materials. Polyphosphazenes are unique compared to more common polymer

systems which have carbon, oxygen, nitrogen, or silicon in the backbone. Polyphosphazenes are inorganic main-chain polymers with an alternating phosphorus-nitrogen (P-N) backbone and two substituents attached to each phosphorus atom. The top of Figure 1 shows the general structure for a side-chain liquid crystal polymer: a polymer backbone with a side chain comprised of a flexible spacer and a rigid rod (mesogen). The remaining structures in Figure 1 are examples of side-chain liquid crystal polyphosphazenes. The middle structure represents a phosphazene mixed-substituent "copolymer" which contains a mesogen side chain along with a non-mesogen (R) side chain. Phosphazene homopolymers (Figure 1, bottom) contain two identical substituents attached to phosphorus.

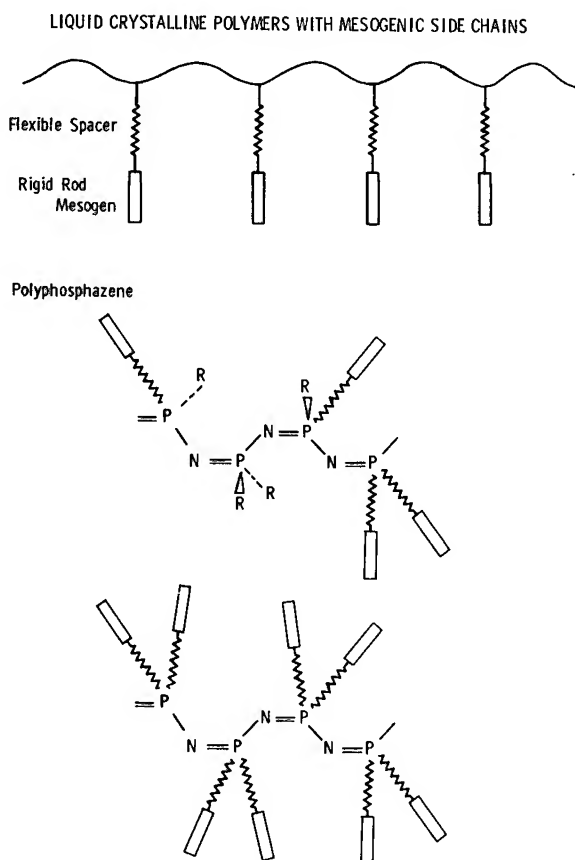


Figure 1. General structures for side-chain liquid crystal polymers.

SYNTHESIS

The most commonly used approach for the synthesis of polyphosphazenes involves the ring-opening polymerization of hexachlorocyclotriphosphazene (trimer) to give open-chain high molecular weight poly(dichlorophosphazene), followed by a substitution process to yield hydrolytically stable poly(organophosphazenes) (3,4). This synthetic procedure is outlined in Figure 2. By using highly purified trimer and by carefully monitoring the polymerization process, one can obtain soluble poly(dichlorophosphazene) and avoid branching and crosslinking which leads to "inorganic rubber". The second step is also crucial: poly(dichlorophosphazene) is a reactive inorganic macromolecule, and chlorine substitution is required in order to prepare hydrolytically stable polymers. This also allows one to prepare a variety of different polymers. Shown in Figure 2 are three general examples consisting of alkoxy, aryloxy, and amino substituted phosphazene polymers. Commercial poly(fluoroalkoxy- and aryloxyphosphazene) elastomers have been prepared, where two or more nucleophiles are used in the substitution process (5). Phosphazene high-performance elastomers are used in both commercial and military applications. Polyphosphazenes which contain organometallic and bioactive side chains have been reported (5).

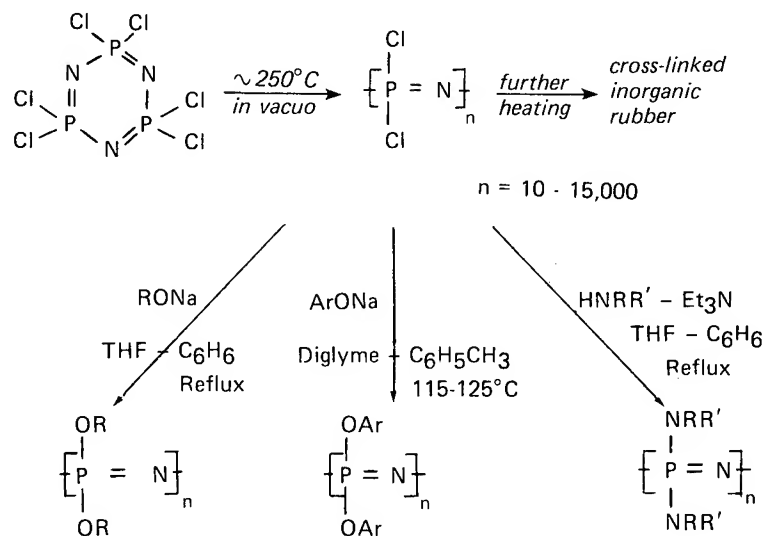


Figure 2. Synthesis of poly(dichlorophosphazene) and poly(organophosphazenes).

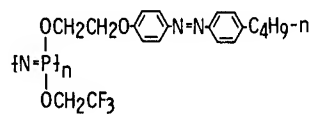
The two-step synthesis process shown in Figure 2 affords several options in preparing new side-chain liquid crystal polymers. The polymerization allows one to vary the molecular weight and molecular weight distribution (MWD), and potentially change the properties of the liquid crystalline state. Most of the polyphosphazenes reported in the literature, including the examples in this paper, are derived from the bulk uncatalyzed polymerization process: this generally produces high molecular weight polymer (one million or greater) with a broad MWD. Catalyst systems have been developed which give lower molecular weight poly(dichlorophosphazene) with a narrower MWD (5). One catalyst, boron trichloride, has also been used to prepare poly(dichlorophosphazene) with narrow molecular weight fractions ranging from approximately ten thousand to over one million (6). Thus, it may be possible to prepare side-chain liquid crystal polyphosphazenes with a range of molecular weights and change the properties of the mesophase.

Polyphosphazene synthesis provides additional possibilities for preparing liquid crystal polymers with different properties. As noted above, the substitution process (Figure 2) enables one to synthesize a wide variety of polymers. The phosphazene inorganic backbone is a highly flexible polymer chain: glass transition temperatures can vary from below room temperature to above 100°C, depending on the size and the nature of the group attached to the P-N backbone (4,5). Crystallinity can be altered by preparing mixed-substituent "copolymers" (7). Thus, changing the side-chain substituent affords the potential for varying both the nature and temperature range of the mesophase.

For the synthesis of side chain liquid crystal polyphosphazenes, the most important examples in Figure 2 are the alkoxy- and aryloxy-polymers. Mesophase behavior has been noted with simple side chains, such as trifluoroethoxy and aryloxy side chains (8). This mesophase behavior is not conventional liquid crystal order, but polymers which exist in a conformationally disordered state (9).

LIQUID CRYSTAL POLYPHOSPHAZENES

In Figure 3 are examples of side-chain liquid crystal polymers prepared in our laboratory. Phosphazene copolymer (I) contains a mixture of an azophenoxy mesogen and a trifluoroethoxy non-mesogen (ca. 1.3:0.7) side chain (10). Structure I in Figure 3 represents only one of the possible repeat units, since the distribution of substituents is presumed to be random (see Figure 1). Although the copolymer (I) contained only slightly greater than one mesogen side chain per repeat unit, it formed a reversible thermotropic liquid crystal phase upon cooling from the

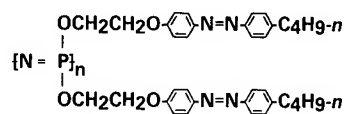


$$\frac{\text{Mesogen}}{\text{Alkoxide}} \sim \frac{1.3}{0.7}$$

ISOTROPIC STATE ABOVE 175°C

MESOPHASE BETWEEN 123°C AND 175°C

I

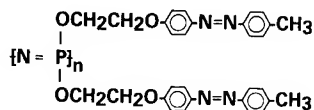


ISOTROPIC STATE ABOVE 171°C

MESOPHASE BETWEEN 144°C AND 171°C

FOCAL CONIC FAN TEXTURE - SMECTIC PHASE

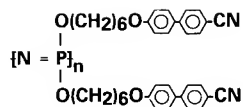
II



ISOTROPIC STATE ABOVE 135°C

MESOPHASE BETWEEN 114°C AND 135°C

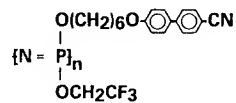
III



ISOTROPIC STATE ABOVE 110°C

MESOPHASE BETWEEN 58°C AND 110°C

IV



$$\frac{\text{Mesogen}}{\text{Alkoxide}} \sim \frac{1.6}{0.4}$$

ISOTROPIC STATE ABOVE 92°C

MESOPHASE BETWEEN 57°C AND 92°C

V

Figure 3. Side-chain liquid crystal polyphosphazenes.

isotropic state above 175°C . Analysis by differential scanning calorimetry (DSC) showed no prominent first-order transitions for either the heating or cooling cycles, which may be due to the irregular distribution of substituents along the P-N backbone. Microscopic analysis in the liquid crystalline region is indicative of a nematic texture (Figure 4). The polymer crystallized upon cooling to room temperature. However, upon heating above 120°C , the polymer transformed back into a mesophase which persisted up to the final clearing point.

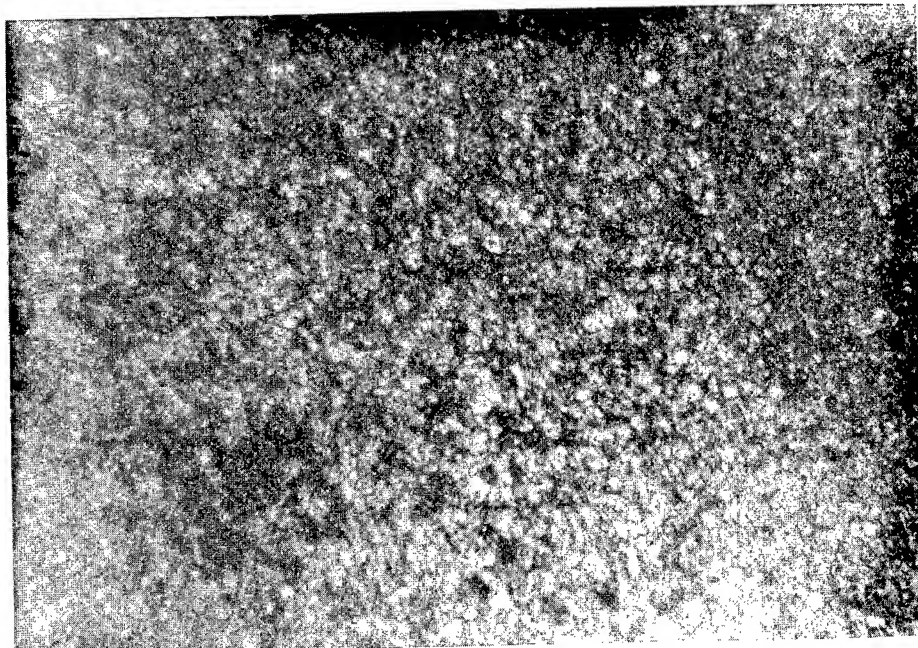


Figure 4. Optical micrograph of polymer I at 128°C . Crossed polarizers. Magnification 192X. Reference 10.

A different situation exists for phosphazene polymer II, shown in Figure 3, containing solely the *n*-butyl phenylazophenoxyethoxy side chain. On cooling from the isotropic state (above 171°C) a viscous mesophase is first observed which shows the characteristic focal conic texture typical of low molar mass smectic A phases. A photomicrograph of the smectic A phase at 159°C is shown in Figure 5 (11). As the temperature falls below 159°C , further texture changes occur. Below 145°C , the polymer crystallizes. DSC analysis of II, taken on the cooling cycle from the isotropic melt, is shown in Figure 6. As with polymer I, the thermo-optical behavior is reversible.



Figure 5. Optical micrograph of polymer II at 159°C. Crossed polarizers. Magnification 200X. Reference 11.

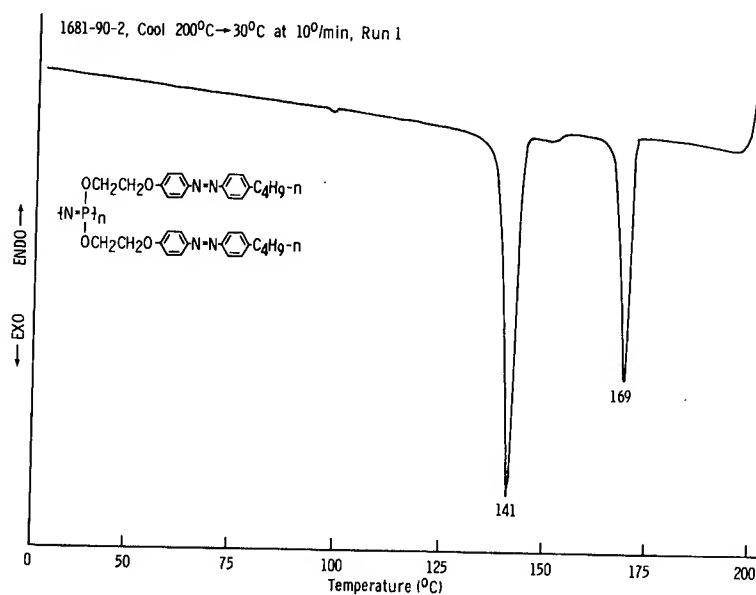
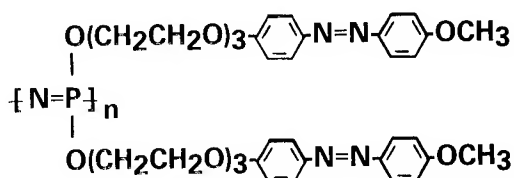


Figure 6. DSC cooling curve of polymer II at 10°C/min.

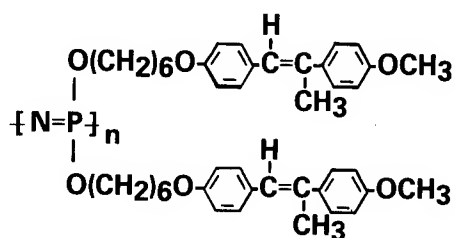
Additional examples of polyphosphazenes (III-V) are shown in Figure 3. These polymers also show reversible thermotropic liquid crystalline behavior. Mesogens include para-substituted phenylazophenols and para-cyanobiphenol. Both two-carbon and six-carbon spacer compounds have been prepared. Polymers have been studied by differential scanning calorimetry (DSC) and hot-stage optical microscopy. The thermal transition behavior is summarized in Figure 3. The nature of the mesophase(s), mesophase temperature range, and isotropization point are dependent on the spacer length and type of mesogen. Polymers are both film- and fiber-forming materials and are soluble in common solvents.

A different series of side-chain liquid crystalline polyphosphazenes and the analogous cyclic trimers were reported by Allcock and Kim (12,13). The side chain consisted of aromatic azo mesogenic units attached to the phosphorus atoms through flexible oligomeric ethyleneoxy spacer units. Polymers were studied by DSC and optical microscopy. For polymer VI, the transition from the crystalline to the liquid crystalline state occurred at 118°C, and this phase formed an isotropic fluid at 127°C. On cooling, the transition from the isotropic to the liquid crystalline state occurred at 126°C, with a final crystallization at 94°C. Polarized optical microscopy of the mesophase at 121°C showed a typical nematic pattern. This study showed that the terminal para substituent on the mesogen, the length of the spacer unit, and the cyclic or long-chain polymer character of the phosphazene all played a role in determining whether or not the material was liquid crystalline and also played a role in determining the nature of the liquid crystalline state.



VI

The study by Percec, Tomazos and Willingham (14) looked at the influence of polymer backbone flexibility on the phase transition temperatures of side-chain liquid crystalline polymethacrylate, polyacrylate, polymethylsiloxane and polyphosphazene containing a stilbene side chain. Upon cooling from the isotropic state, polymer VII displayed a monotropic nematic mesophase between 106 and 64°C. In this study, the polymers with the more rigid backbones displayed enantiotropic liquid crystalline behavior, whereas the polymers with the flexible backbones, including the siloxane and the polyphosphazene, displayed monotropic nematic mesophases. The examples in this study demonstrated how kinetically controlled side chain crystallization influences the thermodynamically controlled mesophase through the flexibility of the polymer backbone.

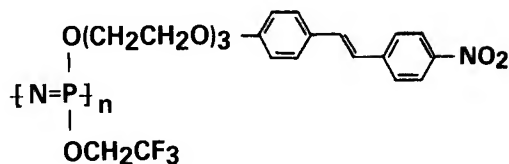


VII

NONLINEAR OPTICAL POLYMERS

Part of the rationale for investigating liquid crystal polyphosphazenes is for nonlinear optical (NLO) applications. In general, for NLO activity, polymers must either contain noncentrosymmetric side chains (eg., side-chain liquid crystal polymers) or highly delocalized (conjugated) backbones (15). Alternatively, polymers can be doped with low molar mass compounds to obtain NLO activity.

Recently, a mixed-substituent polyphosphazene (polymer VIII) was synthesized and the second-order NLO properties were investigated (16). The nitrostilbene/trifluoroethoxy ratio was approximately 36:64. Due to the low glass transition temperature of VII ($T_g = 25^\circ\text{C}$), the second harmonic signal decayed to zero within a few minutes.⁸ However, polymer VII is a prototype which offers many opportunities for further tailoring the molecular structure of polyphosphazenes to generate an optimum combination of NLO and physical properties.



VIII

Another recent report (17) involves the evaluation of phosphazenes with nonmesogen side chains that exhibit NLO activity. Modification of the intrinsic optical response of the P-N chain was achieved through substitution of side groups including chloro, anilino, dimethylamino, and trifluoroethoxy. Cyclic phosphazenes were also included in this study. Results suggest that phosphazenes possess inherent NLO activity which can be enhanced by the suitable selection of substituent groups.

MOLECULAR MODELING

In addition to polymer synthesis and characterization, we have recently begun to evaluate the potential of computer-aided molecular modeling as a tool for designing new liquid crystalline polymers and NLO materials based on phosphazene chemistry. This approach to the design of chemical compounds has been developed and utilized over the past 10 to 15 years by the pharmaceutical industry for the design of bioactive molecules. As a result of the rapid growth in computing power available to the average research laboratory, it has been possible to increase both the sophistication and the scope of applications of the modeling techniques. Commercial software products intended specifically to model polymers have been recently introduced by several companies, and numerous research establishments and manufacturers have initiated molecular modeling efforts to assist in the development of new materials.

Tripos Associates' SybylTM software, including the polymer modeling package, is being used at MTL in conjunction with an Evans and Sutherland PS390 graphics terminal connected to the Polymer Research Branch VAX cluster local area computer network. The molecular modeling system can predict the optimal (lowest energy) conformation of molecules and can calculate intrinsic electronic properties, such as polarizability and dipole moment. The system also performs calculations to predict certain bulk properties of crystalline polymers, such as tensile modulus, based on small polymer fragments and arrays extrapolated to infinite dimensions. Molecular mechanics calculations are most frequently used to determine lowest energy conformations of molecules, but the system also provides access to semi-empirical molecular orbital calculations when it is necessary to consider electronic effects in a system. The Sybyl software at MTL also includes a Quantitative Structure-Property Relationships (QSPR) module which can be used to correlate molecular structure with physical properties on a statistical basis. This is often the fastest way to make predictions on the properties of new molecular designs.

Preliminary results at MTL show a good correlation between experimental structure data obtained by X-ray diffraction analysis of one of our liquid crystalline phosphazene polymers and the conformation predicted by the modeling program. We anticipate that modeling of the polymers and chromophores will provide insight as to how the liquid crystalline properties and NLO response of these materials can be controlled by varying the molecular structures.

CONCLUSION

At the present time, only a few side-chain liquid crystal polyphosphazenes have been synthesized and investigated. Based on the synthesis process outlined in Figure 2, opportunities exist to prepare a wide variety of new polyphosphazenes with potential for NLO applications such as high speed switching devices, sensors, and eye protection against low power lasers. Alternative synthetic approaches, such as side-chain modification of polyphosphazenes prepared by the thermal decomposition of N-silylphosphoranimines (18), may provide even further options for preparing new NLO active polyphosphazenes.

ACKNOWLEDGEMENTS

We wish to acknowledge the collaboration with Dr. Claudine Noel, E.S.P.C.I., Paris, France; Dr. Edwin Atkins, Department of Physics, University of Bristol, United Kingdom; Dr. Robert Lenz, Polymer Science and Engineering Department, University of Massachusetts, Amherst, MA; Dr. Geoff Lindsay, Naval Weapons Center, China Lake, CA.

REFERENCES

1. Ulrich, D. R.; Mol. Cryst. Liq. Cryst. 1988, 160, 1.
2. Side-Chain Liquid Crystal Polymers; McArdle, C. B., Ed.; Blackie, Chapman and Hall, New York, 1989.
3. Allcock, H. R. Phosphorus-Nitrogen Compounds, Academic: New York, 1972.
4. Singler, R. E.; Schneider, N. S.; Hagnauer, G. L. Polym. Eng. Sci. 1975, 15, 321.
5. Inorganic and Organometallic Polymers; Zeldin, M.; Wynne, K. J.; Allcock, H. R., Eds.; ACS Symposium Series No. 360, American Chemical Society, Washington, DC, 1989; Chapters 19-21.
6. Sennett, M. S., U.S. Patent 4 867 957, 1989.
7. Beres, J. J.; Schneider, N. S.; Desper, C. R.; Singler, R. E. Macromolecules 1979, 12, 566.
8. Schneider, N. S.; Desper, C. R.; Beres, J. J. In Liquid Crystalline Order in Polymers; Blumstein, A., Ed.; Academic: New York, 1978; p 299.
9. Wunderlich, B.; Grebowicz, J. Adv. Polym. Sci. 1984, 60/61, 1.
10. Singler, R. E.; Willingham, R. A.; Lenz, R. W.; Furukawa, A.; Finkelmann, H. Macromolecules 1987, 20, 1727.
11. Singler, R.E.; Willingham, R.A.; Noel, C.; Friedrich, C.; Bosio, L.; Atkins, E.D.T.; Lenz, R.W. Polym. Prepr. (Amer. Chem. Soc. Div. Polym. Chem.) 1989, 30(2), 491.

12. Kim, C.; Allcock, H. R. Macromolecules 1987, 20, 1726.
13. Allcock, H. R.; Kim, C. Macromolecules 1989, 22, 2596.
14. Percec, V.; Tomazos, D.; Willingham, R. A. Polym. Bull. Berlin 1989, 22, 199.
15. Nonlinear Optical Properties of Organic Molecules and Crystals; Chemla, D. S.; Zyss, J., Eds.; Academic, New York, 1987.
16. Dembek, A. A.; Kim, C.; Allcock, H. R.; Devine, R. L. S.; Steier, W. H.; Spangler, C. W. Chem. Mater. 1990, in press.
17. Exharos, G. J.; Samuels, W. D. Presented at the Meeting of the Materials Research Society, Boston, MA, December 1989; paper S2.4
18. Neilson, R. H.; Wisian-Neilson, P. Chem. Rev. 1988, 88, 541.

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

BINARY ANTIDOTES FOR ORGANOPHOSPHATE CHEMICAL WARFARE AGENTS (U)

RUTHANN M. SMEJKAL*, Dr., HAIM LEADER, Dr., A. DAVID WOLFE, Dr.,
RICHARD K. GORDON, Dr., AND PETER K. CHIANG, Dr.
DIVISION OF BIOCHEMISTRY, WALTER REED ARMY INSTITUTE OF RESEARCH
WASHINGTON, DC 20307-5100

At cholinergic muscarinic receptors throughout the body, the presynaptic release of acetylcholine results in the binding of this neurotransmitter to muscarinic receptors, causing an end effect, such as muscle contraction or release of secretagogues from stimulated cells. Unbound acetylcholine is quickly degraded by acetylcholinesterase to maintain a crucial equilibrium between the bound and unbound transmitter. Inactivation of acetylcholinesterase results in restimulation of the cells before recovery can occur and can lead to incapacitation or death.

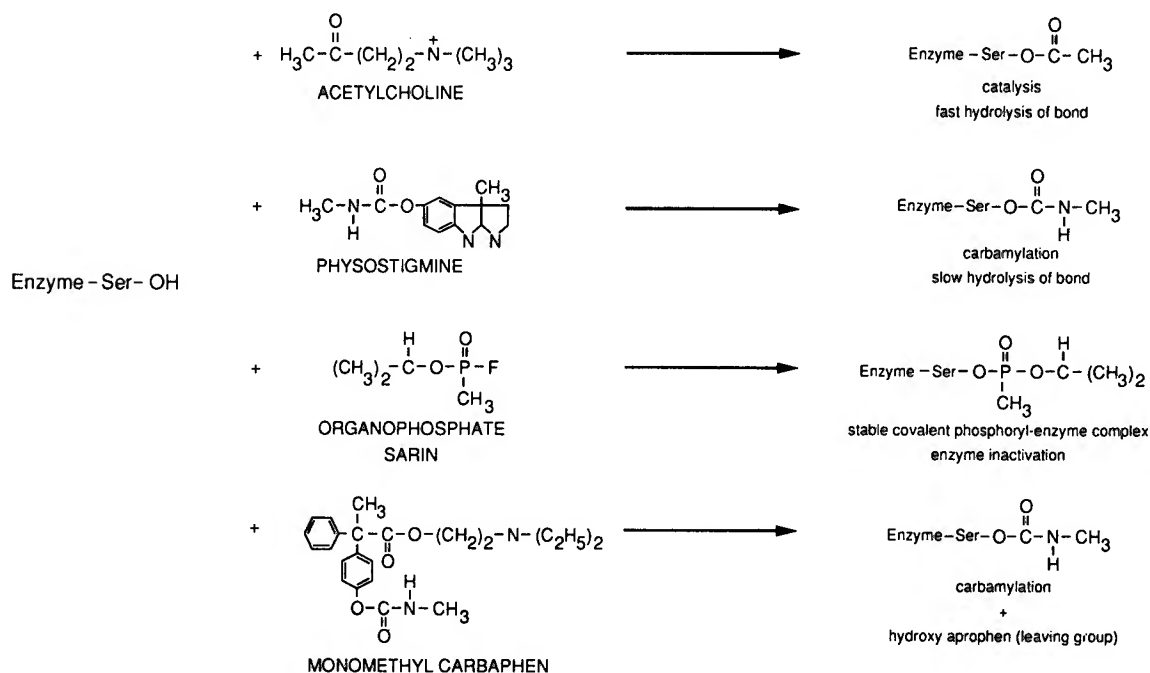
The chemistry behind these interactions is shown in Scheme 1. Acetylcholine acetylates a serine moiety at the active-site of cholinesterases with choline as the leaving group. This newly-formed bond is quickly hydrolyzed, releasing the active enzyme.¹ Carbamates, such as physostigmine or pyridostigmine, reversibly carbamylate the active-site of cholinesterases, permitting reactivation of the enzyme through decarbamylation.^{2,3} Organophosphates also act at the active-site serine, forming a stable covalent phosphoryl-enzyme complex, thereby irreversibly inactivating the enzyme, causing an accumulation of acetylcholine and preventing cell recovery.^{4,5} Pretreatment with carbamates offers prophylactic protection against organophosphate poisoning because these drugs chemically sequester cholinesterases from the action of organophosphates, while permitting recovery of enzyme activity through decarbamylation.^{3,6-9}

In treating casualties of organophosphate poisoning, various therapeutic strategies may be employed. Aprophen (N,N-diethylaminoethyl 2,2-diphenylpropionate) is a potent anticholinergic and antispasmodic agent possessing a wide number of distinct pharmacological actions, including both antimuscarinic and noncompetitive nicotinic antagonist activities.¹⁰⁻¹⁶ The potent antimuscarinic and antinicotinic effects of aprophen make it a potential drug of choice in the therapy of poisoning by organophosphate/anticholinesterase agents.¹⁶⁻¹⁸ The combination of

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

aprophen and a carbamate significantly improves the protection afforded by either drug alone against the physiological and behavioral symptoms of organophosphate poisoning, even if post-poisoning treatment were omitted.¹⁸

MECHANISMS OF INTERACTION WITH CHOLINESTERASE ENZYMES

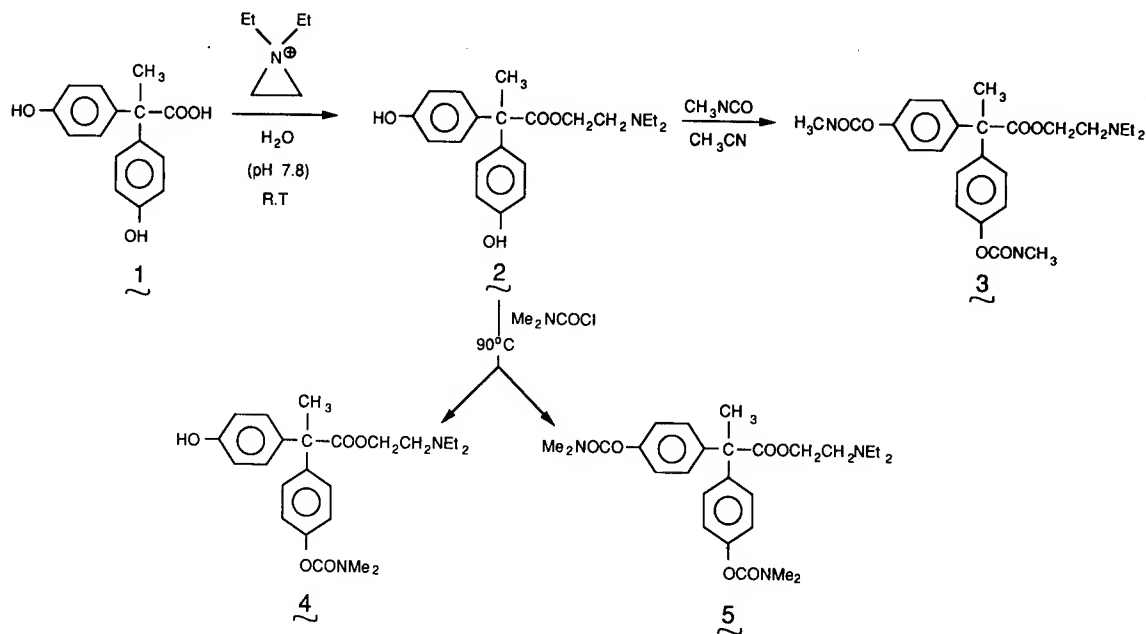


Scheme 1. Mechanisms of Interaction with Cholinesterases

In view of the remarkable antidotal efficacy obtained by the combination of physostigmine and aprophen, we have synthesized a series of aprophen analogs with carbamyl substitutions on one or both of the phenyl rings. The rationale behind the design of such compounds is that they might act as 'binary' (bifunctional) drugs, exhibiting most of the therapeutic characteristics of aprophen, along with the ability to protect prophylactically by chemically sequestering cholinesterases. In addition, the presence of two cholinergic recognition sites per molecule might cause higher local concentrations of the drug to be available at the post-synaptic sites.

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

Chemistry. Carbamyl-aprophen compounds, which we have termed 'carbaphens', were synthesized with para-substitution on both phenyl rings of the aprophen molecule (series 1) or on only one phenyl ring of aprophen (series 2) by N-monomethylcarbamyl or N,N-dimethylcarbamyl groups. The synthetic pathway of the dicarbamyl-aprophen compounds is outlined in Scheme 2.

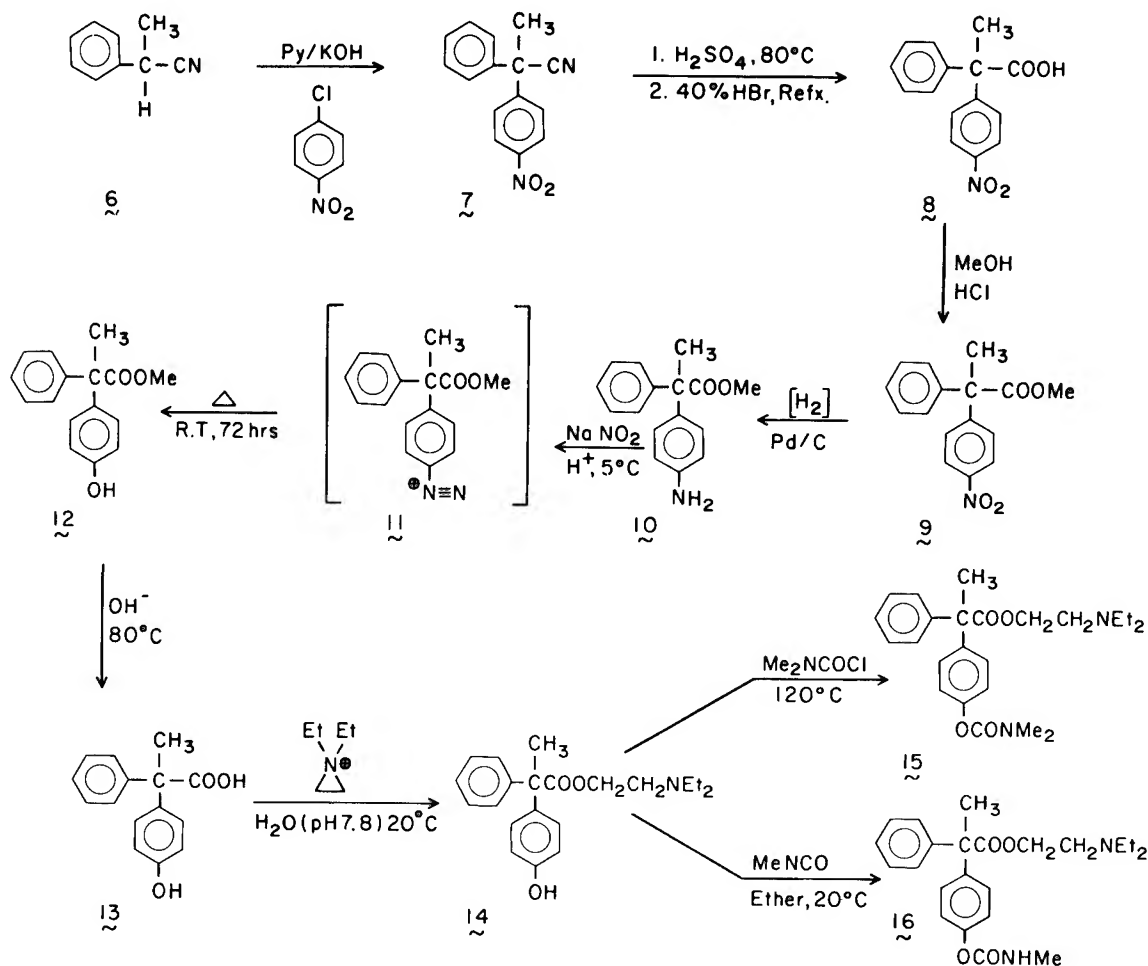


Scheme 2: Synthetic Pathway of the Disubstituted Aprophen Derivatives.

The conversion of the key intermediate, 2,2-bis-(4-hydroxyphenyl) propionic acid 1, to its diethylaminoethyl ester 2 did not proceed satisfactorily using known conventional methods (acylchloride, transesterification, etc.). The ester 2 (dihydroxy aprophen) was obtained in 86% yield, however, after a NaHCO₃ solution of 1 was mixed and stirred at room temperature with an aqueous solution of N,N-diethylaziridinium chloride, in the presence of ethyl acetate. Treatment of dihydroxy aprophen 2 with an excess of methylisocyanate in acetylnitrile yielded the bismonomethylcarbamate 3 (bismonomethyl carbaphen), which was purified by column chromatography. Treatment of dihydroxy aprophen 2 with excess dimethylcarbamylchloride at 90° C gave a

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

mixture of the monocarbamate 4 (dimethylhydroxy carbaphen) and the dicarbamate 5 (bisdimethyl carbaphen), which were separated by column chromatography.



Scheme 3: Synthetic Pathway of the Monosubstituted Aprophen Derivatives.

The synthetic pathway for the mono-carbamyl aprophen derivatives is outlined in Scheme 3. α-Methylbenzyl-cyanide 6 was converted to the p-nitrophenyl substituted nitrile 7, and hydrolyzed via the amide to the appropriate propionic acid 8. The methyl ester 9 was hydrogenated to the amine 10, which was diazotized to 11 and decomposed at room temperature to give the hydroxy methyl ester derivative 12. Alkaline hydrolysis of the methyl ester 12 afforded the hydroxy acid 13 which was esterified to the hydroxy aprophen derivative 14. Dimethylaminocarbamyl aprophen 15

(dimethyl carbaphen) was obtained by reacting hydroxy aprophen 14 with excess dimethylcarbamyldichloride at 120° C for 2 hrs. Monomethylcarbamyldichloride aprophen 16 (monomethyl carbaphen) was obtained by a Na-catalyzed addition reaction of methylisocyanate to 14 in ether solution. All the new compounds were purified by column chromatography, and their structural elucidation and purity were confirmed by ¹H NMR, CI-MS and TLC.

Biological Activities. These compounds were tested for antimuscarinic activity (the ability to compete with acetylcholine at muscarinic receptors) and anticholinesterase potency (the ability to inactivate a cholinesterase by carbamylation, as shown in Scheme 1). Antimuscarinic activity was assessed by three biological assays. The analogs were tested for their ability to antagonize the acetylcholine-induced contraction of guinea pig ileum, carbachol-stimulated release of α -amylase from pancreatic acini, and the binding of [³H]N-methylscopolamine to the muscarinic receptor sites of cerebral cortex membranes.¹⁴⁻¹⁶ The anticholinesterase activity of the compounds were determined colorimetrically by the method of Ellman,¹⁷ as described by Main, et al.¹⁸ by testing their effects on the activities of human serum butyrylcholinesterase, electric eel acetylcholinesterase and fetal bovine serum acetylcholinesterase.¹⁹ Preliminary in vivo studies were performed to test the efficacy of a carbaphen against an organophosphate.

Antimuscarinic activities: Table I shows that the most potent antimuscarinic analog was obtained by substituting a monomethylcarbamate moiety at the para-position of one phenyl ring of aprophen, monomethyl carbaphen (16). The pA₂ for monomethyl carbaphen was 7.3 ± 0.1 (K_B = 5.0 x 10⁻⁸ M) when tested on the acetylcholine-induced contraction of guinea pig ileum, which has predominantly the m₂ muscarinic receptor subtype.^{16,16,25-28} It inhibited the carbachol-induced release of α -amylase from pancreatic acinar cells, which have an m₃ muscarinic receptor subtype,²⁶⁻³⁰ with a K_i of 2.0 ± 0.9 x 10⁻⁹ M. With respect to the inhibition of [³H]NMS binding to cerebral cortex membranes, which contain predominantly the m₁ receptor subtype,^{28,31,32} the K_i was 2.6 ± 1.5 x 10⁻⁸ M. Overall, the antimuscarinic profile of monomethyl carbaphen was very similar to that of aprophen (Table I), with the former showing more specificity for the pancreatic m₃ muscarinic receptor subtype over the m₁ subtype of the cortex.

While the pA₂ value for the inhibition of ileum contraction by dimethyl carbaphen (15), was about the same as that of monomethyl carbaphen (16), the inhibition of α -amylase release and [³H]NMS binding were each one order of magnitude lower (Table I). Substitution with a para-dimethylcarbamate on one phenyl ring of aprophen and a para-hydroxyl group on the other [dimethylhydroxy carbaphen (4)] was not well tolerated. The antimuscarinic activities of 4 were one to two orders of magnitude lower than those of dimethyl carbaphen. Substitution with a

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

TABLE I

ANTIMUSCARINIC ACTIVITY OF CARBAPHEN ANALOGS AND STANDARD COMPOUNDS*

Compound	Guinea Pig		Inhibition of		Cerebral cortex [³ H]NMS binding K _i (M)
	Ileum Contraction pA ₂	[K _B (M)]	Pancreatic Acini α-amylase Release K _i (M)		
Monomethyl carbaphen, <u>16</u>	7.3 ± 0.1	5.0 x 10 ⁻⁸	2.0 ± 0.9 x 10 ⁻⁹		2.6 ± 1.5 x 10 ⁻⁸
Dimethyl carbaphen, <u>15</u>	7.1 ± 0.1	8.3 x 10 ⁻⁸	4.2 ± 0.9 x 10 ⁻⁸		1.5 ± 0.5 x 10 ⁻⁷
Hydroxy apropphen, <u>14</u>	7.2 ± 0.3	7.0 x 10 ⁻⁸	5.6 ± 1.2 x 10 ⁻⁹		1.0 ± 0.5 x 10 ⁻⁷
Dihydroxy apropphen, <u>2</u>	5.7 ± 0.1	2.0 x 10 ⁻⁶	4.9 ± 0.8 x 10 ⁻⁶		2.1 ± 0.2 x 10 ⁻⁶
Bismonomethyl carbaphen, <u>3</u>	not active		not active		1.3 ± 1.1 x 10 ⁻⁵
Dimethyl- hydroxy carbaphen, <u>4</u>	5.5 ± 0.2	3.2 x 10 ⁻⁶	2.4 ± 0.8 x 10 ⁻⁶		3.4 ± 0.3 x 10 ⁻⁶
Bisdimethyl carbaphen, <u>5</u>	not active		not active		marginal activity
Apropphen	8.5 ± 0.1	3.1 x 10 ⁻⁹	1.7 ± 0.7 x 10 ⁻⁹		5.1 ± 1.0 x 10 ⁻⁸
Atropine	8.7 ± 0.1	2.0 x 10 ⁻⁹	1.6 ± 1.1 x 10 ⁻⁹		2.4 ± 0.7 x 10 ⁻⁹
Pirenzepine	4.3 ± 0.4	5.0 x 10 ⁻⁵	1.2 ± 0.2 x 10 ⁻⁷		1.7 ± 0.7 x 10 ⁻⁷

*Each inhibition constant represents the mean of four to five independent experiments ± standard errors of the mean.

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

para-monomethyl carbamate moiety on both phenyl rings of aprophen [bis-monomethyl carbaphen(3)] yielded a compound that was a poor inhibitor of [^3H]NMS binding and was inactive in both the ileum contraction and α -amylase release assays. Bisdimethyl carbaphen (5), a compound with para-dimethylcarbamate substitution on both phenyl rings of aprophen, was inactive in all three assays. In comparing the inhibition constants of 15, 2 and 4, it was evident that these carbaphen analogs did not exhibit any muscarinic subtype specificity.

Interestingly, the mono-para-hydroxy-substituted aprophen, hydroxy aprophen (14), which is both the synthetic precursor of mono and dimethyl carbaphen and the decarbamylated end product, was almost as potent an antimuscarinic as monomethyl carbaphen and showed the most preference for the m_3 muscarinic receptor subtype of the pancreas over the m_1 subtype of the cerebral cortex (about 18-fold). However, the addition of a second para-hydroxyl group to aprophen [dihydroxy aprophen (2)] led to a drastic decrease in antimuscarinic activity with inhibition constants in the range of 10^{-6} M.

These structure-activity relationship data suggest that one of the phenyl rings of aprophen must remain unsubstituted in order to preserve antimuscarinic activity, while the substitution of a carbamate or hydroxyl function in the para-position of the other phenyl ring is well tolerated. This is evident from the poor antimuscarinic potency of dihydroxy aprophen (2) and dimethylhydroxy carbaphen (4), and the lack of activity of bismonomethyl carbaphen (3) and bisdimethyl carbaphen (5). In contrast, compounds 14, 15, and 16, which have one unsubstituted phenyl ring, are potent antimuscarinic agents.

Anticholinesterase Activity: When tested for their ability to inactivate human serum butyrylcholinesterase, i.e. carbamylating ability, the five carbaphen analogs, monomethyl carbaphen (16), dimethyl carbaphen (15), bismonomethyl carbaphen (3), dimethylhydroxy carbaphen (4) and bisdimethyl carbaphen (5) showed a potent ability to carbamylate butyrylcholinesterase (Table II). The inactivation constants (K_I values) were on the order of 10^{-6} M, similar to those of pyridostigmine and physostigmine, which are clinically useful carbamates. It should be noted that aprophen, itself, was not subject to hydrolysis by human serum butyrylcholinesterase.³³

In contrast to the potent inactivation of butyrylcholinesterase, the carbaphen analogs showed marginal time-dependent inactivation (about 20 % inhibition at drug concentrations greater than 10^{-4} M) against acetylcholinesterase purified from fetal bovine serum²⁴ or obtained from electric eel (data not shown). These results are consistent with the inhibition, observed using steady-state kinetics, produced by aprophen^{34,35} and other aromatic/neurotoxic agents,³⁶⁻³⁸ which preferentially inhibit butyrylcholinesterase over acetylcholinesterase.

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

TABLE II

ANTI-BUTYRYLCHOLINESTERASE (HUMAN) ACTIVITY OF CARBAPHEN ANALOGS*
AND STANDARD COMPOUNDS

Carbamates	inactivation constant, K_I (M)	bimolecular rate constant, k_3 ¹ (liter/mole min)
Monomethyl carbaphen, <u>16</u>	$1.4 \pm 0.6 \times 10^{-6}$	$8.5 \pm 4.2 \times 10^5$
Dimethyl carbaphen, <u>15</u>	$1.3 \pm 1.0 \times 10^{-6}$	$6.6 \pm 3.5 \times 10^5$
Bismonomethyl carbaphen, <u>3</u>	$7.2 \pm 5.6 \times 10^{-6}$	$2.9 \pm 0.2 \times 10^5$
Dimethylhydroxy carbaphen, <u>4</u>	$0.8 \pm 0.5 \times 10^{-6}$	$3.0 \pm 2.3 \times 10^5$
Bisdimethyl carbaphen, <u>5</u>	$4.6 \pm 0.3 \times 10^{-6}$	$1.4 \pm 0.8 \times 10^4$
Pyridostigmine	$3.5 \pm 1.3 \times 10^{-6}$	$3.4 \pm 1.1 \times 10^4$
Physostigmine	$5.0 \pm 2.4 \times 10^{-6}$	$7.0 \pm 0.9 \times 10^4$

*Each inhibition constant represents the mean of four to five independent experiments \pm standard errors of the mean.

Similar results with additional compounds have led to the conclusion that butyrylcholinesterase, in comparison with acetylcholinesterase, contains a unique hydrophobic site which partially determines the substrate and inhibitor patterns of butyrylcholinesterase.^{39,40}

In Vivo Anti-Organophosphate Activity: For preliminary in vivo protection studies, a highly toxic anti-cholinesterase organophosphate, 7-(methylethoxyphosphinyloxy)-1-methylquinolinium iodide (MEPQ), was used. MEPQ inhibits acetylcholinesterase at an extremely high rate with a 1:1 stoichiometry, and has an LD₅₀ in guinea pigs of approximately 25 μ g/kg, permitting the use of very small volumes.⁴¹ With no pretreatment, guinea pigs die within about 15 min after intramuscular injection of 2 LD₅₀ of MEPQ and within 10 min after intramuscular injection of 4 LD₅₀ of MEPQ. Following a 30 min pretreatment with 10 mg/kg of monomethyl carbaphen (16) administered intramuscularly (a dose sufficient to inhibit

cholinesterase in vivo by about 60%), guinea pigs survived a challenge of 2 LD₅₀ of MEPQ. Of five pretreated animals challenged with 4 LD₅₀ of MEPQ, two animals died after about 15 min, one died after about 45 min and two survived the crisis and were still alive and feebly moving around at the end of the day. Thus, monomethyl carbaphen demonstrates protection against multiple LD₅₀ doses of a potent organophosphate, even in the absence of post-treatment with supportive drugs.

Conclusion. The magnitude of the inhibition constants of the carbaphen analogs against butyrylcholinesterase show that, in contrast to the steric requirements of the muscarinic receptors, bulky phenyl ring substitutions of aprophen can be tolerated by butyrylcholinesterase. However, these analogs are only marginally active against acetylcholinesterase. It is important to note that both 15 and 16, after carbamylating butyrylcholinesterase, produce as an end product hydroxy aprophen (14), itself a potent antimuscarinic (Table I). Thus, the carbaphen analogs, monomethyl carbaphen (16) and dimethyl carbaphen (15), are binary prototype drugs that are potent both as antimuscarinic agents and as carbamates. Preliminary in vivo protection studies have demonstrated that the monomethyl carbaphen will carbamylate serum butyrylcholinesterase and offers protection against multiple LD₅₀ doses of the organophosphate MEPQ. Therefore, these compounds show promise as prophylactic/therapeutic antidotes for organophosphate poisoning, as well as being candidates for combination treatment modalities.

REFERENCES

1. Rosenberry, T. L. *Proceed. Natl. Acad. Sci. USA* 72, 3834 (1975).
2. Green, A. L. *Biochem. Pharmacol.* 32, 1717 (1983).
3. Aldridge, W. N. in *Enzyme Inhibitors as Drugs*, Ed. M. Sandler, MacMillan, New York, 115 (1980).
4. Koelle, G. B. in *The Pharmacological Basis of Therapeutics*, 4th ed., Eds. L. S. Goodman and A. Gilman, MacMillan, New York, 442 (1970).
5. Korolkovas, A.; Burekhalter, J. H. in *Essentials of Medicinal Chemistry*, John Wiley and Sons, New York, 212 (1976).
6. Berry, W. K.; Davies, D. R. *Biochem. Pharmacol.* 19, 927 (1970).
7. Hayashi, E.; Okudaira, H.; Yamada, S. *Toxicol. Appl. Pharmacol.* 48, 111 (1979).
8. Lennox, W. J.; Harris, L. W.; Talbot, B. G.; Anderson, D. R. *Life Sci.* 37, 793 (1985).
9. McGee, J.; Brezenoff, H. E. *Life Sci.* 41, 65 (1987).
10. Mashkovsky, M. D.; Liberman, S. S. *Farmakol. Toksikol.* 20, 354 (1957).
11. Volkova, Z. V. *Farmakol. Toksikol.* 22, 348 (1959).

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

12. Brown, N. D.; Smejkal, R. M.; Breuer, E.; Doctor, B. P.; Chiang, P. K. *J. Pharmaceut. Sci.* 77, 145 (1988).
13. Nakazato, Y.; Oleshansky, M. A.; Chiang, P. K. *Arch. Int. Pharmacodyn.* 293, 209 (1988).
14. Beach, J. E.; Smallridge, R. C.; Chiang, P. K.; Fein, H. G. *J. Pharmacol. Exp. Ther.* 246, 548 (1988).
15. Carroll, F. I.; Abraham, P.; Parham, K.; Griffith, R. C.; Ahmad, A.; Richard, M. M.; Padilla, F. N.; Witkin, J. M.; Chiang, P. K. *J. Med. Chem.* 30, 805 (1987).
16. Witkin, J. M.; Gordon, R. K.; Chiang, P. K. *J. Pharmacol. Exp. Ther.* 242, 796 (1987).
17. Leadbeater, L.; Inns, R. H.; Rylands, J. M. *Fund. Appl. Toxicol.* 5, S225 (1985).
18. Leadbeater, L.; D'Mello, G. D. *Proc. 2nd Int. Symp. Protection Against Chemical Warfare Agents, Stockholm, Sweden.* pp. 335 (1986).
19. Pankaskie, M. C.; Kachur, J. F.; Itoh, T.; Gordon, R. K.; Chiang, P. K. *J. Med. Chem.* 28, 1117 (1985).
20. Gordon, R. K.; Chiang, P. K. *J. Pharmacol. Exp. Ther.* 236, 85 (1986).
21. Ahmad, A.; Gordon, R. K.; Chiang, P. K. *FEBS Lett.* 214, 285 (1987).
22. Ellman, G. L.; Courtney, K. D.; Anders, V., Jr.; Featherstone, R. M. *Biochem. Pharmacol.* 7, 88 (1961).
23. Main, A. R.; Soucie, W. G.; Buxton, I. L.; Arinc, E. *Biochem. J.* 143, 733 (1974).
24. Ralston, J. S.; Rush, R. S.; Doctor, B. P.; Wolfe, A. D. *J. Biol. Chem.* 260, 4312 (1985).
25. Peralta, E. G.; Ashkenazi, A.; Windslow, J. W.; Smith, D. H.; Ramachandran, J.; Capon, D. J. *EMBO J.* 6, 3923 (1987).
26. Gordon, R. K.; Breuer, E.; Padilla, F.; Smejkal, R. M.; Chiang, P. K. *Molec. Pharmac.* 36, 766 (1989).
27. Barlow, R. B.; Weston-Smith, P. *Br. J. Pharmacol. Sci.* 85, 437 (1985).
28. Doods H. N.; Mathy, M.-J.; Davidesko, D.; Van Charldorp, K. J.; De Jong, A.; Van Zwieten, P. A. *J. Pharmacol. Exp. Ther.* 242, 257 (1987).
29. Hootman, S. R.; Picado-Leonard, T. M.; Burnham, D. B. *J. Biol. Chem.* 260, 4186 (1985).
30. Korc, M.; Ackerman, M. S.; Roeske, W. R. *J. Pharmacol. Exp. Ther.* 240, 118 (1987).
31. McKinney, M.; Richelson, E. *Ann. Rev. Pharmacol. Toxicol.* 24, 121 (1985).
32. McKinney, M.; Stenstrom, S.; Richelson, E. *Mol. Pharmacol.* 27, 223 (1985).
33. Aarbakke, J.; Muira, G. A.; Brown, N. D.; Gray, R. R.; Gordon, R.

UNCLASSIFIED

SMEJKAL, LEADER, WOLFE, GORDON AND CHIANG

- K.; Doctor, B. P.; Chiang, P. K. J. Pharm. Pharmacol. 38, 928 (1986).
34. Verdier, J. S.; Wolfe, A. D. Biochem. Pharm. 5, 1605 (1986).
 35. Marquis, J. Neurotoxicol. 6, 261 (1985).
 36. Augustinsson, K.-B. Acta Physiol. Scand. 15 (Suppl 52), 1 (1948).
 37. Perkinson, E.; Ruckert, R.; Da Vanzo, J. P. Proc. Soc. Exp. Biol. Med. 131, 685 (1969).
 38. Maayani, S.; Weinstein, H.; Ben-Zvi, N.; Sokolovsky, M. Biochem. Pharm. 23, 1263(1974).
 39. Kabachnik, M. I.; Brestkin, A. P.; Godovikov, N. N.; Michelson, M. J.; Rozengart, E. V.; Rozengart, V. I. Pharmacol. Revs. 22, 355 (1970).
 40. Augustinsson, K.-B. Biochim. Biophys. Acta 128, 351 (1966).
 41. Levy, D. Ashani, Y. Biochem. Pharm. 35, 1079 (1986).

UNCLASSIFIED

The Effects of Electron Beam Processing on High Electron Mobility Transistors (U)

*Doran D. Smith, Dr., William D. Braddock, Mr, and
Mitra Dutta, Dr.

U. S. Army Electronics Technology and Devices Laboratory
Fort Monmouth, New Jersey 07703-5000

In recent years the technology of electron-beam lithography for submicron and ultrasubmicron device fabrication has come into extensive use. As device size shrinks the possibility of electrons achieving ballistic transport throughout the device, and hence maintaining phase coherence through out the device becomes a possibility. Devices that operate on this principle include electron waveguide devices that work by confining the 2D EG formed at a heterojunction interface to a quasi 1D channel^{1,2}. However, for these devices to be successful it is necessary to maintain high mobility material throughout processing so that the electrons can transit a device without losing phase coherence. Several patterning techniques are employed for fabricating ultrasubmicron devices. Ion implantation is used for direct modification of heterostructures, and electron beam and x-ray lithography are used to pattern resists which are then used as a mask for pattern transfer to the underlying wafer. In this paper we describe the experimental determination of the effect of e-beam irradiation on the transport properties of a GaAs-AlGaAs modulation doped heterostructure.

Figure 1 shows the structure that was grown by molecular beam epitaxy (MBE); on an undoped GaAs substrate, as described previously³, were deposited a 1 μ m undoped GaAs buffer layer, a 200 Å undoped Al_{0.3}Ga_{0.7}As layer, a 350 Å Al_{0.3}Ga_{0.7}As layer doped with $2.8 \times 10^{17}/\text{cm}^3$ silicon, and a 250 Å GaAs layer doped with $1.0 \times 10^{18}/\text{cm}^3$ silicon. This results in the interface that creates the

250Å	GaAs	Si= 1.0×10^{18}
350Å	Al _{0.3} Ga _{0.7} As	Si= 2.8×10^{17}
200Å	Al _{0.3} Ga _{0.7} As	

1.0μ	GaAs Buffer	Undoped

S.I.	GaAs	Substrate Undoped

Fig. 1 The modulation doped heterostructure used.

2D EG channel 800 Å below the surface. The as-grown material was fabricated into Hall bars using optical lithography and wet chemical etching. Ohmic contacts were made to the contact pads of the Hall bar using indium balls which were annealed in hydrogen at 425 °C for five minutes. The mobility and number density of each Hall bar were then individually determined.

Measurements were performed with the sample directly immersed in liquid helium at 4.2 K. A magnetic field was applied normal to the face of the sample and a current density of $5 \times 10^{-3} \text{ A/m}$ was used to excite the Hall bar. The Shubnikov-de Haas (SdH) oscillations in ρ_{xx} and ρ_{xy} were characteristic of a 2D EG, where the ρ_{xx} minimum in the quantum Hall region went to zero, indicating no parallel conduction in the as-grown material. The 2D EG number densities, n_s , as determined from the periodicity of the SdH oscillations between 0.7 and 2.0 T were $3\text{--}4 \times 10^{11} \text{ /cm}^2$. The resistivity ρ_{xx} at zero magnetic field was obtained from $\rho_{xx}(B=0) = [V_{xx}(B=0)/L]/[I/W]$. V_{xx} is the voltage between adjacent side arms, L is the spacing between adjacent side arms (1 mm), W is the width of the Hall bar (200 μm), and I is the current flowing in the Hall bar. The zero magnetic field drift mobility was obtained from $\mu = 1/[\rho_{xx}(B=0)en_s]$. The mobility values for all samples were $5\text{--}6 \times 10^5 \text{ cm}^2/\text{V}\cdot\text{sec}$.

After their initial characterization, the samples were electron irradiated in a Cambridge Instruments EBMF 10.5 at electron energies between 2.5 - 20 keV. The electron beam dose chosen for this study was that relevant for an isolated feature in the commonly used electron beam resist polymethylmethacrylate (PMMA), which may be as high as 1600 $\mu\text{C/cm}^2$. To set an upper limit on the damage that can be expected, 1600 $\mu\text{C/cm}^2$ was chosen.

Figure 2(a) shows ρ_{xx} for an as-grown, before 10 keV electron irradiation. Below 0.7 T the SdH oscillations at 4.2K are too small to be observable, and by 3 to 4 T the quantum Hall effect dominates. Figure 2(b) shows ρ_{xx} after 10 keV electron irradiation. The electron damage produced parallel conduction and damped out the oscillations at low magnetic field. The 2D EG concentration was virtually unchanged in all samples.

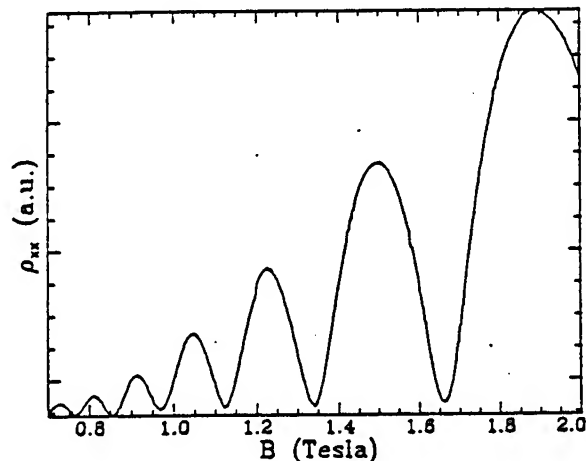


Fig. 2a ρ_{xx} for the 10 keV sample before irradiation.

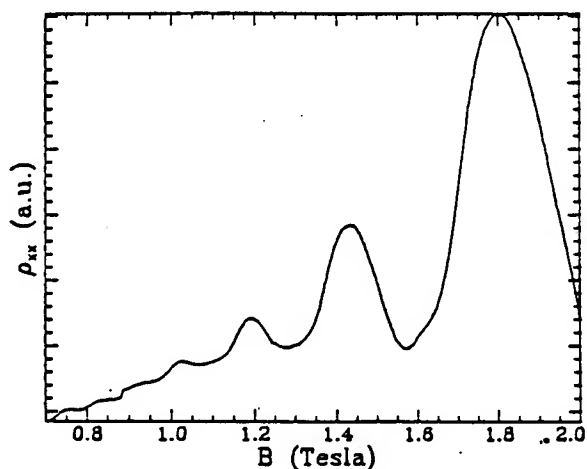
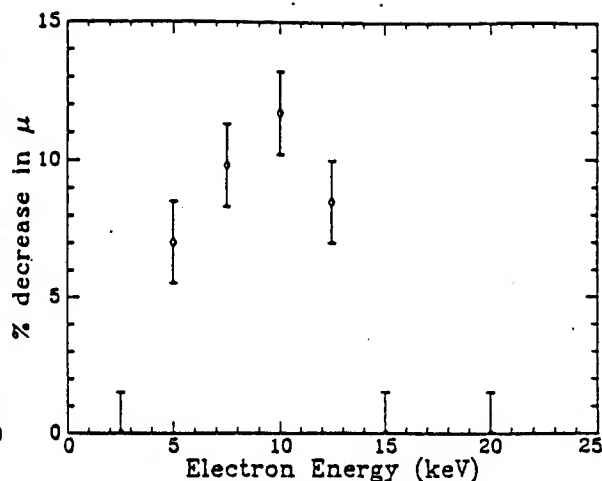
Fig. 2b ρ_{xx} after irradiation.

Fig. 3 2D EG drift mobility degradation after irradiation.

Figure 3 shows the percentage decrease in mobility after electron irradiation as a function of electron energy. Mobility was not degraded at 2.5, 15, and 20 keV, but showed a decrease at 5, 7.5, 10, and 12.5 keV, with the maximum reduction in mobility at 10 keV. To understand the shape of the mobility degradation curve it is necessary to understand electron energy deposition rate as a function of depth into the heterostructure.

The range⁴ (depth of penetration into the sample) of electrons in a material is predicted by integration of Eq. (1)^{5,6},

$$\frac{dE}{dx} = \frac{2\pi e^4 n}{E} \ln \frac{1.16E}{I}, \quad (1)$$

where E is electron energy, x is distance, n is the number of electrons in the stopping medium per cubic centimeter, and I is the geometric mean of all the ionization and excitation potentials of the absorbing atom⁷ (approximately Z times the Rydberg constant). Equation (1) can be solved by separation and direct integration in terms of the logarithm integral⁸,

$$x = \frac{1}{a b^2} [Ei(2 \ln(b E_0)) - Ei(2 \ln(b E))] \quad (2)$$

where

$$a = \frac{2\pi e^4 \rho Z}{(1.6 \times 10^{-6}) m_n A},$$

$$b = \frac{1.16}{I}.$$

E_0 is the initial energy, E is the energy at actual path length x , m_n is the neutron mass, and $E_i()$ is tabulated⁹. Equation (2) is plotted in fig. 4 for the energies used in this study.

Charged defects created in proportion to dE/dx can explain the energy dependence of the mobility degradation, as illustrated by comparing figures 3 and 4. There was no measurable damage to the crystal at 2.5 keV, which suggest that the damage at 2.5 keV, if any, is below the threshold for damage detection with this method. As the electron energy increases to 5.0, 7.5, and 10.0 keV so does the damage to the crystal, and it is sufficiently close to the 2D EG so as to cause an increasing amount of mobility degradation. Further increases in electron energy cause dE/dx in the region of the 2D EG to continue to decrease at a slower rate; however, dE/dx becomes large at significantly more remote distances. On the high energy side dE/dx became low enough in the region of the 2D EG, and the high damage region was sufficiently remote from the 2D EG, so as not to have a measurable effect on the mobility.

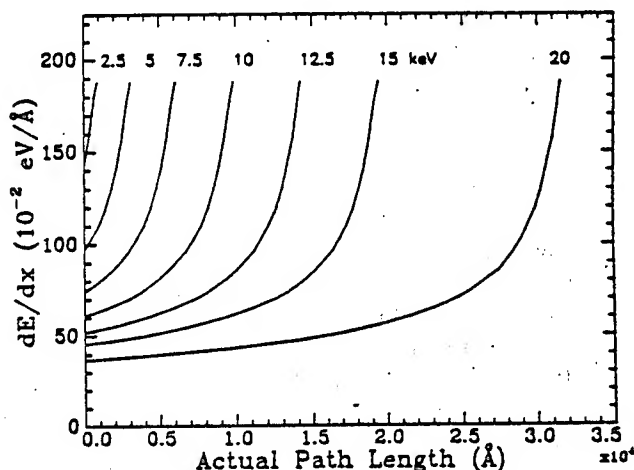


Fig. 4 Electron energy loss rate vs. actual path length for several initial energies.

Neutral defects can also degrade mobility; but fail to fit the energy dependence of dE/dx . For a neutral defect (short range interaction) to scatter an electron in the 2D EG, the wavefunction of the electron must overlap the defect site, requiring the defect to be located very close to the heterojunction interface. Therefore, the shape of the mobility degradation curve should be proportional to dE/dx at the heterojunction interface only; however, it is not.

The previous discussion assumed that charged defects were created in proportion to dE/dx without mentioning the source of the defects. To dislocate an As atom from its respective sublattice, creating a vacancy

and an interstitial (Frenkel pair) requires an incident electron of approximately 250keV, representing a threshold for Frenkel pair creation of approximately 10eV¹⁰. Ga is mobile well below room temperature, so Ga vacancy related defects are not observed in this study¹⁰. Therefore, the defects created in this work are not Frenkel pairs. We propose that the defects are created by modifying already existing defects. In the process of modifying one defect type to another the magnitude of its charge state is being increased. This experiment is unable to determine what type of defect is being observed, but one possibility is hydrogen¹¹, which is 75 times lighter than As, making energy transfer from the incident electron beam 75 times more efficient. Other possibilities include a defect with a mass of Ga or As, with a 100-200meV threshold for conversion.

In actual device fabrication, the wafer will be coated with e-beam resist. When electron energy loss in the resist is taken into account the percent mobility degradation curve will shift to higher energies and perhaps widen. For the typical 1/4 micron gate length HEMT the e-beam resist thickness would be 0.5 to 1 micron thick. The density of PMMA is 1.18 gm/cm³ (other e-beam resists are similar). The combination of density and thickness is sufficient, so there will be significant energy loss in the resist for the energies studied here. However, for the case of ultrasubmicron devices, where the resist thickness is no more than 1000 Å, there will be negligible energy loss in the e-beam resist.

This work preformed in collaboration with Tobin Fink of New Jersey Institute of Technology, Newark, New Jersey 07102.

References

1. G. Timp, H. U. Baranger, P. deVegvar, J. E. Cunningham, R.E. Howard, R. Behringer, and P.M. Mankiewicz, *Phys. Rev. Lett.* 60, 2081 (1988)
2. Fernando Sols, M. Macucci, U. Ravaioli, and Karl Hess, *Appl. Phys. Lett.* 54, 350 (1989).
3. Doran. D. Smith, T. Fink, W. D. Braddock, and Mary-Lloyd Saunders in The Second Workshop on Radiation-Induced and/or Process-Related Electrically Active Defects in Semiconductor-Insulator Systems, edited by A. Reisman (Microelectronics Center of North Carolina, Research Triangle Park, NC, 1989) p. 316.
4. S.V. Starodubtsev and A. M. Romanov, The Passage of Charged Particles Through Matter, (U. S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, Springfield, VA, 1962), p. 196.
5. *ibid*, p 178.
6. The range of an electron at these energies is 50% less than the actual path length the electron travels before stopping due to straggling (a few large angle scattering events). The integration of Eq. (1) should predict the actual path length, not range. However, there is a wealth of experimental data that demonstrates theory over estimates dE/dx in this low energy range, see for example ref. 7 p 624, and references listed there.
7. Robley D. Evans, The Atomic Nucleus, (McGraw-Hill Book Company, Inc., New York, 1955) p. 580.
8. I. S. Gradshteyn and I.M. Ryzhik, Table of Integrals, Series, and Products, (Academic Press, New York, 1965), p. 204.
9. Milton Abramowitz and Irene A. Stegun, Handbook of Mathematical Functions, (Dover Publications, New York, 1965), p. 228.
10. D. Pons, P.M. Mooney, and J. C. Bourgoin, *J. Appl. Phys.* 51, 2038 (1980).
11. Yi-Ching Pao, D. Liu, W. S. Lee, and J. S. Harris, *Appl. Phys. Lett.* 48, 1291, (1986).

A Selectively-Contacted Dual Channel
High Electron Mobility Transistor (U)

Doran D. Smith, Dr.

and *Ravi Khanna, Mr.

U. S. Army Electronics Technology and Devices Laboratory,
Fort Monmouth, New Jersey 07703-5000

I. Introduction

Increasing the speed of operation of transistors has benefits that span the range of the electronic equipment that rely on them. Computing and communication ability increase as the maximum operating frequency of transistors increases. A limiting factor in the switching speed of existing field-effect transistors is the transit time determined by the effective gate length. We have grown and fabricated in-house a novel selectively-contacted dual channel high electron mobility transistor (SCDC HEMT), invented at ETDL¹, in which we attempt to circumvent the gate transit delay by using electron transport in the vertical as well as the usual lateral direction.

We have based our design on a DCHEMT, introduced by Iafrate and modeled by Hess et al., in which each channel of a dual-channel device is selectively contacted^{1,2}. In the modeled device, the two channels, which are adjacent to each other, are inserted between a frontside and backside Schottky barrier gate. By reverse biasing the upper gate and forward biasing the lower one, band bending forces electrons from the upper channel into the lower one. This allows conduction through the lower channel only. The reverse effect is obtained by forward biasing the upper gate and reverse biasing the lower one. Since the distance between the upper and lower channel can be made smaller than the effective gate length, the transit time is shorter.

Our initial results on an earlier device were the first report of a SCDC HEMT³. We have since improved device performance using modifications to the initial design and improved processing techniques. A backside gate is not yet employed, in order to facilitate processing. Because of our unique ability to contact each channel separately, a single device may be operated as an inverter for computing purposes. The selective contacts allow a study of charge transfer from one channel to the other under action of a central gate, in a manner

not possible using standard three-terminal HEMT's. We describe in detail how the device is fabricated and characterized, how magnetotransport measurements are used to gain insight into internal charge transfer processes in the device, and the implications this has for future devices.

II. Detailed device description and fabrication

The device has three Schottky barrier gates (GD, GS, and G) and four ohmic contacts (S1, S2, D1, and D2) (fig. 1). There are two main requirements necessary for selective contact: 1) The metals from the upper channel's ohmic contacts, S1 and D1, must be prevented from penetrating into the lower channel 2) The outer ohmic contacts, S2 and D2, must contact only the lower channel. The epilayer structure of our device consists of a twin quantum well formed by two 30nm undoped GaAs layers separated by a 10nm AlGaAs barrier, the purpose of which is to provide some degree of electrical isolation between the channels. The barrier width of 5nm in the previous design provided insufficient isolation. Two n-AlGaAs layers doped with $1 \times 10^{18}/\text{cm}^3$ Si provide electrons to form two-dimensional electron gas (2D EG) channels at the upper and lower interfaces of the quantum well. In our current design, requirement 1) is met by using non-alloyed delta-doped contacts^{4,5}. In these contacts, a suitable metal is deposited onto a very heavily doped Si layer at the epilayer surface. The metal forms an ohmic contact without being alloyed, insuring that the lower quantum well is not penetrated. In our current design requirement 2) is met alloying ohmic metal at a high temperature. This causes the metal to diffuse deep enough to contact the lower channel, but also shorts the upper and lower channels together. Selective contact is achieved by reverse biasing the isolation gates (GD and GS) so that the upper channel next to each outer (deeper) ohmic contact is depleted. This forces current flow from the outer ohmic contacts through the lower channel, and thus effectively allows selective contact of S2 and D2 to the lower channel only. Once the bias of the isolation gates is set, the bias on the gate G controls device operation.

Device fabrication began with mesa isolation. Contact areas for the outer (deeper) source and drain were formed by depositing Au/Ge/Ni/Ag/Au (60nm/30nm/10nm/100nm/100nm) and by annealing at 550°C for 10 seconds. The delta-doped contacts for the upper source and drain were formed from gold. All gates were formed simultaneously by depositing Ti/Au (25nm/200nm) in a recess made with a pH-7 selective etch. A single resist layer was used to define the gate recess and gate lift-off metallization. Etch depth was verified by monitoring the current between the upper source and drain. After gate lift-off, 100 nm of Si₃N₄ passivation was deposited by plasma enhanced chemical vapor deposition (PECVD) at 300°C. Contact pad windows were etched in

the Si₃N₄ using buffered HF, and, finally Ti/Au (5 nm/100 nm) contact pads were deposited on the Si₃N₄ using electron-beam evaporation.

III. Electrical experiments

To test our understanding of the device operation the device was biased for signal inversion as shown in fig. 1. Reverse biasing the isolation gates was not necessary since the upper channel was already depleted due to a deep gate recess. Under these conditions conduction into S2 and D2 can be through the lower channel only. S1 was set to -0.5V and S2 was set to 0.5V. The output voltage was taken from D1 across a 10Kohm load. When the input voltage V_{in} is negative the output should be close to the potential of S2, since the upper channel from S1 to the output D1 will be depleted due to the reversed bias input gate. When V_{in} is increased positively, electrons are expected to transfer from the lower channel to the upper one, and thus bring the output closer to the potential of S1.

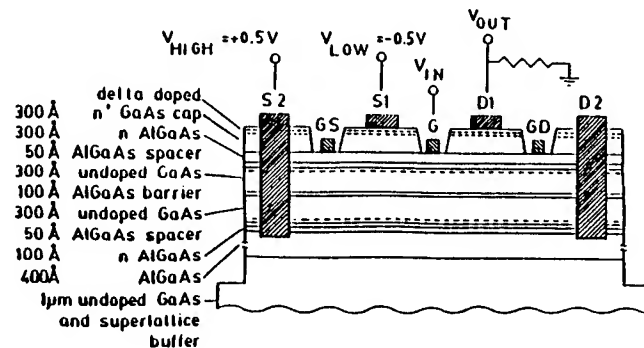


Fig. 1 SCDC HEMT in inverter configuration.

The device was electrically characterized at 300K and 77K. Transconductance at 300K showed a distinctive double-peak behavior with a peak of 60mS at 0.25V and a smaller peak of 15mS at -0.74V (fig. 2a). The input-output transfer curve is shown in fig. 2b.

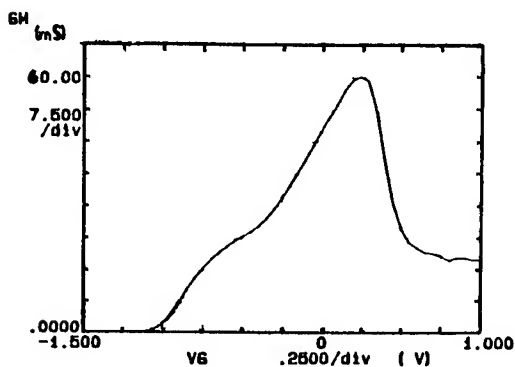


Fig. 2a Transconductance at 300K.

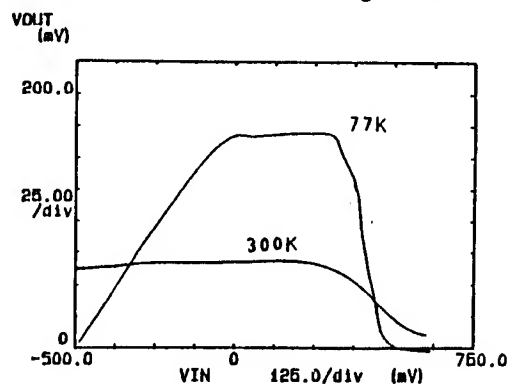


Fig. 2b Input-output transfer curve for a 10 Kohm load.

300K operation is now possible and yields an output voltage swing of -60mV for an input voltage swing of 0.35V. The output voltage swing at 77K is -175mV for an input voltage swing of only 200mV. This is a great improvement over the previous device³, and to the authors knowledge is the first inverting behavior obtained by this method. An unusual feature of fig. 2b is the presence of both positive and negative transfer characteristics within the bias range.

IV. Magnetotransport Experiments

To learn more about how charge redistributes itself under the gate G as a function of gate bias, a series of gated Hall bars were fabricated and magnetotransport experiments performed on them (fig.3). It is possible, as will be described below, to perform an experiment which yields a signal, part of which, is manifestly related to the charge density in each channel. Therefore, it is possible to determine the charge density in each channel as a function of gate bias.

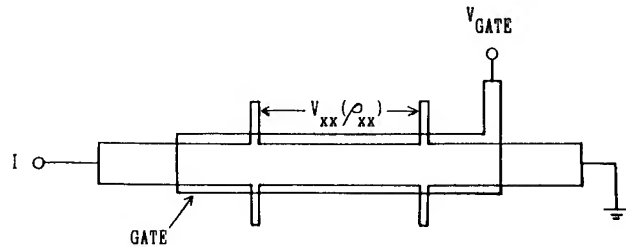


Fig. 3 Hall bars used in the magnetotransport studies. See Text for symbol description.

The epilayer structure of the gated Hall bars was similar to that used to fabricate the SCDC HEMT devices. Samples with two different quantum well widths were studied. One sample was symmetric, containing two quantum wells of 22.2nm wide separated by a barrier of 7.0nm. The second sample was asymmetric, containing one quantum well of 22.2nm and a second of 8.1nm. The selection of the well widths was chosen so that: 1) the total well width in the first sample was large enough to guarantee two separate 2D EG channels, and 2) the asymmetry of the well widths of the second sample would guarantee different ground state energies in the two channels.

All magnetotransport measurements were performed at 4.2K by direct immersion in liquid helium. A current of 1 or 10 microamps was passed down the length of the Hall bar, and a magnetic field of 0-9 Tesla was applied normal to the sample surface. ρ_{xx} is proportional to V_{xx} shown in fig. 3. ρ_{xx} is given by equ. 1 as a function of magnetic field⁶.

$$\rho(B) = \rho_0 + [1 + (w_c \tau)^2] \frac{\sigma_s}{\sigma_0^2} \quad (1)$$

where $\rho_0 = \frac{1}{\sigma_0}$

and
$$\sigma_s = - \frac{\sigma_0}{1 + (w_c \tau)^2} 2 \frac{(w_c \tau)^2}{1 + (w_c \tau)^2} \frac{2\pi^2 kT}{hw_c} \operatorname{csch} \left[\frac{2\pi^2 kT}{hw_c} \right] \cos \left[\frac{2\pi E_f}{hw_c} \right] e^{-\pi/w_c \tau}$$

where σ_0 is the Drude conductivity, σ_s contains the SdH oscillations, w_c is the cyclotron frequency, E_f is the Fermi level, and τ is the quantum mechanical scattering time. $E_f = n_s \pi \hbar^2 / m$, where n_s is the charge density in a 2D EG layer, and m is the effective mass of the electron. $w_c = eB/mc$, where e is the electron charge, and B is the magnetic field. The oscillatory part of equ. 1 is sensitive only to the charge in the 2D EG. Hence measurement of ρ_{xx} is an excellent technique to separate the effects of 2D charges from 3D charges (which might, for example, reside in the n-AlGaAs), since 3D charges will not cause oscillations in ρ_{xx} under the conditions described here.

Typical data as a function of magnetic field taken on the symmetric sample is shown in fig. 4. There are two features present not described in equ. 1. The positive slope of ρ_{xx} (positive magnetoresistance) in the 0.5-2 Tesla region is caused by both 2D EG channels conducting. This feature is unimportant for the work described here, and will not be discussed further. The oscillatory part of equ. 1 is periodic in $1/B$. While fig. 4 shows a strong component periodic in $1/B$, it also contains some features that appear aperiodic in $1/B$. This is because two subbands, one in each quantum well, are conducting. Hence, there are really two oscillations periodic in $1/B$. To see this more clearly, the non-oscillatory

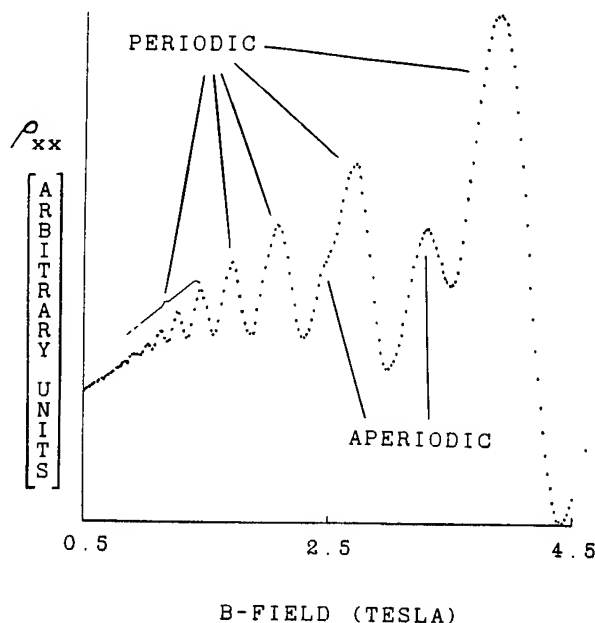


Fig. 4 ρ_{xx} data as a function of magnetic field for the symmetric device described in fig. 3. I is 1 μ amp, the gate voltage is -0.145 V

background is subtracted from the data. The resulting oscillatory part is replotted with $1/B$ as the x-axis. The plot is then Fourier transformed and the result shown in fig. 5. The presence of two peaks in fig. 5 means there are two 2D EG's in the sample. By using the cos argument of equ. 1, the location of the peaks are converted into a charge density in each channel as shown in fig. 5.

Carrying out the procedure outlined above for many different gate voltages applied to both samples results in a determination of charge density in each channel as a function of gate voltage. The result is shown in fig. 6 for the symmetric sample.

Both samples display similar behavior as a function of gate bias. One channel's charge density decreases while the other channel remains constant. For both samples we assume the channel which remains constant to be the lower channel since it is unaffected by the gate voltage. Note also, that as the upper channel depletes its charge, the lower channel does not accumulate the charge.

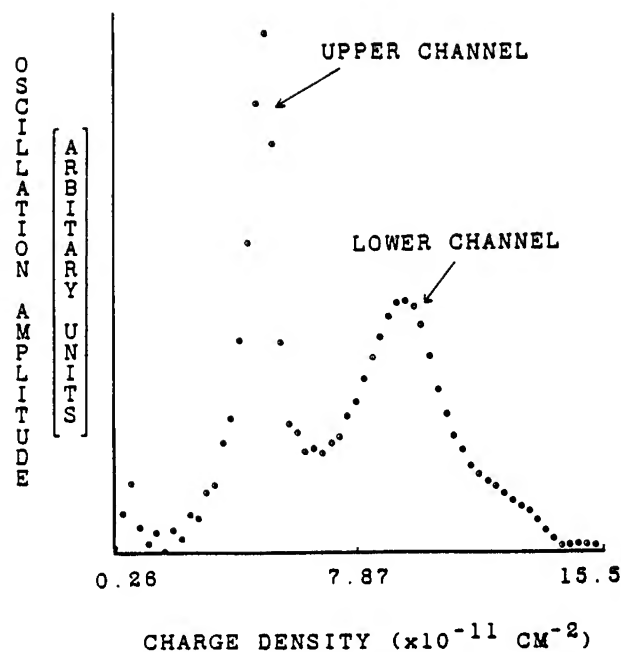


Fig. 5 Fourier transform of the data in figure 5 after removal of the background and transformation to $1/B$ space.

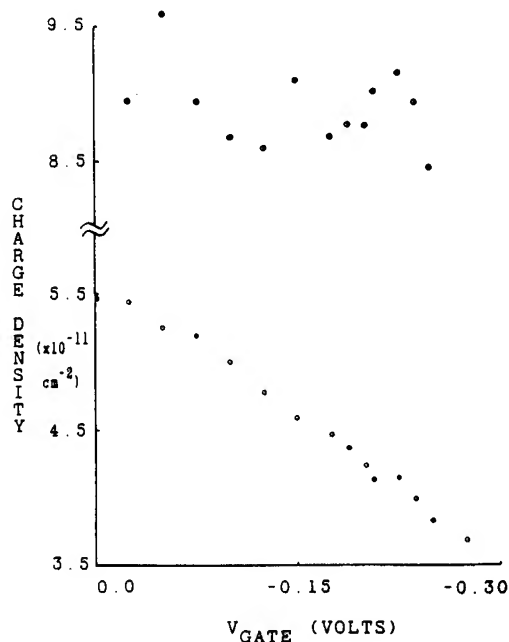


Fig. 6 2D EG charge density as a function of gate voltage for the symmetric sample.

V. Discussion

The magnetotransport measurements show that as the gate bias becomes more negative, the lower channel can deplete only after the majority of charge in the upper channel has been depleted. In addition, a bias condition does not exist when the lower channel is depleted and the upper channel is not, as is predicted for a SCDC HEMT with a backside gate^{1,2}. This is not surprising since a backside gate has been omitted from the present device in order to facilitate fabrication. In spite of this limitation, significant signal inversion occurs, indicating that it is possible to maintain each channel at a different potential at a value high enough to be useful.

This interpretation is further supported by the transconductance and signal inversion measurements. The double-peak behavior of the transconductance (fig. 2a) indicates the presence of two electrically separated channels, since poorly separated channels would appear as a single large channel with a single-peak curve. In addition, the smaller peak occurs in reverse bias. Since the lower channel is farther away from the gate, its transconductance is expected to be less, and to be subject to modulation only after depletion of the upper channel. This would occur only under reverse bias. In the signal inversion measurements the effectiveness of the mid-well AlGaAs barrier will be increased due to reduction of the thermal energy of channel electrons. This results in improved isolation between the channels and a corresponding increase in the potential difference, thus producing the improved inverter behavior observed.

VII Conclusion.

The ability to selectively contact each channel of a dual channel HEMT has resulted in signal inversion from a single device. The inverter transfer curve is greatly improved from the previous design and is unique in that it exhibits both positive and negative transfer characteristics. Magnetotransport experiments have been used to determine vertical charge density distributions within the dual-channel quantum well, and simple electrical measurements yield useful information on channel isolation and charge distribution because selective contacts were used. The performance of the SCDC HEMT is impressive since a backside gate is not being used. To the authors knowledge the SCDC HEMT represents the only instance which inverter-like behavior has been obtained from a single device. Due to its single component nature, this device has the potential for use as a high speed inverter.

This work was performed in collaboration with M. B. Das of the Center for Electronic Materials and Processing, Pennsylvania State

SMITH,

KHANNA

University, University Park, PA 16802, and supported in part by the National Science Foundation/U.S. Army Electronics Technology and Devices Laboratory Interaction Program under Grant ECS-8718801.

REFERENCES

- 1) G. J. Iafrate, L. C. Poli, T. R. AuCoin, and L. F. Heath, "Dual channel high electron mobility transistor," S.N. 228795, July 29, 1988.
- 2) I. C. Kizilyalli, K. Hess, and G. J. Iafrate, "Electron transfer between adjacent channels simulated by ensemble Monte Carlo methods," J. Appl. Phys., vol. 61, no. 6, pp.2395-2398, Mar. 1987.
- 3) R. Khanna, M. B. Das, Doran D. Smith, G. J. Iafrate and P. G. Newman, "A selectively contacted dual-channel HEMT," IEEE Electron Device Lett., vol. EDL-10, no. 12, pp.531-533, Dec. 1989.
- 4) E. F. Schubert, J. E. Cunningham, W. T. Tsang, and T. H. Chiu, "Delta-doped ohmic contacts to n-GaAs," Appl. Phys. Lett., vol. 49, no. 5, pp. 292-294, Aug. 1986.
- 5) G. Gillman, B. Vinter, E. Barbier, and A. Tardella, "Experimental and theoretical mobility of electrons in delta-doped GaAs," Appl. Phys. Lett., vol. 52, no. 12, pp. 972-974, Mar. 1988.
- 6) K.-K. Choi, D. C. Tsui, and S. C. Palmateer, Phys. Rev. B, 33, 8216 (1986).

SMITH, SANDERS, GALES, AND GROSS

Sulfur Mustard-induced Biochemical Alterations
in Proliferating Human Cells in Culture (U)

*William J. Smith, Dr., Kenneth M. Sanders, SPC,
Yolanda A. Gales, SPC, and Clark L. Gross, Mr.
U.S. Army Medical Research Institute of Chemical Defense
Aberdeen Proving Ground, MD 21010-5425

Sulfur mustard (HD) is an alkylating agent which has been shown to have mutagenic (1,2), cytotoxic(3), and vesicating (4) properties. Its use in combat situations has resulted in lethal, incapacitating, and disfiguring injuries. The principal incapacitating injuries come from the vesicating capacity of HD, i.e., production of skin blisters.

Despite decades of medical research, the mechanism by which HD induces vesication is not known and no effective antidotes are currently available. In 1985, Papirmeister et al. (5) proposed a hypothesis delineating the following events for the generation of HD-induced pathology: HD alkylates DNA, DNA strand breaks occur, poly(ADP-ribose)polymerase is activated and utilizes cellular NAD⁺ as substrate, NAD⁺ is depleted, glycolysis is inhibited, hexose monophosphate shunt is activated, proteases are released, and pathology results. As predicted by this hypothesis, HD has been found to lower NAD⁺ levels in both human skin grafted to nude mice(6) and human leukocytes (7). This depletion of cellular NAD⁺ in human lymphocytes is blocked by treatment with niacinamide, an inhibitor of poly(ADP-ribose) polymerase (8). This protection is accompanied by an amelioration of the histologic demonstration of pathology in lymphocytes (9)

Since the dividing basal epidermal cell in skin appears to be the major target of HD, proliferating human epidermal keratinocytes (HEK) and mitogen (Concanavalin A) stimulated human peripheral blood lymphocytes (PBL) were evaluated as in vitro models of HD-induced skin injury. The studies in this report utilize flow cytometric analyses of PBL and HEK for the toxicologic assessment of HD-induced damage. These techniques allowed a unique series of experiments to be conducted which defined levels of sensitivity in human cells to both the cytotoxic and genotoxic effects of HD. These studies have provided data which are relevant to the development of therapeutic intervention in the human pathology produced by HD.

SMITH, SANDERS, GALES, AND GROSS

METHODS:

Reagents:

The lymphocyte studies used RPMI 1640 containing gentamicin (50 $\mu\text{g/ml}$) for cell washing and RPMI 1640 supplemented with 5% fetal bovine serum for final cell suspension and assay (RPMI-FBS). All tissue culture reagents were from GIBCO, Grand Island, NY. Buoyant density isolations were performed using histopaque, $d=1.077$ (Sigma, St. Louis, MO).

Keratinocyte studies used HEK, Keratinocyte Growth Medium (KGM), trypsin-EDTA, and trypsin neutralizing solution from Clonetics Corporation (San Diego, CA).

HD (2,2'-dichlorodiethyl sulfide) was obtained from the Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD. The following chemicals were purchased from commercial sources: RNase A from Boehringer Mannheim (Indianapolis, IN), propidium iodide (PI), dextran T-500, and Concanavalin A (Con A) from Sigma (St. Louis, MO).

PI was dissolved in RPMI 1640 at 20 $\mu\text{g/ml}$ for use in viability studies and at 250 $\mu\text{g/ml}$ in 5% triton X-100 for DNA staining. RNase (1 mg/ml) was added to the DNA staining solution when needed (10).

Lymphocyte Isolation:

Human PBL were isolated from whole blood obtained by venipuncture from volunteers under an approved human use protocol. One ml of 6% dextran was added per 10 ml of blood to expedite red blood cell sedimentation. After allowing sedimentation for 30 minutes at 37°C, the white cell rich supernate was collected, layered onto histopaque ($d=1.077$), and centrifuged at 450 X g for 20 minutes. The mononuclear cell layer was harvested, washed twice in RPMI, and centrifuged at 250 X g for 10 minutes. The isolated PBL were resuspended in RPMI-FBS for cell counting, and initial cell viability determinations were performed by trypan blue dye exclusion. Con A activation of the cells was accomplished by incubating the PBL for 24-72 hours at 37°C in RPMI-FBS containing Con A at a final concentration of 5 $\mu\text{g/ml}$.

Keratinocyte Culture System:

HEK cultures were maintained in 75 cm^2 flasks in KGM at 37°C in 5% CO_2 . When the HEK cultures became 80% to 90% confluent, cells were removed with trypsin-EDTA. After removal, the cells were placed in a conical tube containing trypsin-neutralizing solution, and centrifuged at 250 X g for 10 minutes. Cells were resuspended in 10 ml of KGM for cell counts and initial viability determinations.

SMITH, SANDERS, GALES, AND GROSS

Microplate Preparation:

PBL were resuspended in RPMI-FBS at 4×10^6 cells/ml, and 50 μ l/well dispensed into 96-well microtest plates (Falcon, Oxnard, CA) to yield a final concentration of 2×10^5 cells/well. RPMI-FBS and alkylating agent were added to the wells to yield a final volume of 200 μ l/well. The plates were vented in a fume hood for 1 hour to exhaust volatile agent. They were then incubated at 37°C for 0-72 hours in 5% CO₂ after which time they were prepared for flow cytometric analysis.

HEK (1×10^5 cells/well) were placed in the wells of a 24-well plate in a total volume of 1 ml. The plates were incubated for 24 hours. The media was discarded, and fresh media containing alkylating agent was added to a final volume of 1 ml. The plates were vented and incubated for 24-96 hours.

Flow Cytometry:

PBL. PI in RPMI was added to each well to yield a final concentration of 6 μ g/ 10^6 cells. After 3 minutes at room temperature, the plates were placed on a Multiple Sample Delivery System (MSDS, Coulter Electronic, Hialeah, FL) attached to an EPICS C flow cytometer (Coulter) for viability analysis. The flow cytometer operates with a 5 watt argon laser generating a 488 nm line at 200 mW. The MSDS permits automatic sampling of all wells of the 96-well plate.

Immediately after conclusion of the viability analysis, 50 μ l of PI in triton X-100 was added to each well, and the plate was returned to the MSDS for cytometric analysis of DNA.

Data analysis was performed using the EASY 88 computer system with cell viability analyzed using the Intgra program and DNA histogram analysis by the Paral program (Coulter).

HEK. To determine HEK viability by propidium iodide (PI), the supernatant media containing the floating dead cells were collected in 12 x 75 mm tubes. Cells were trypsinized from the wells, pooled in the respective tubes containing supernatant cells and trypsin neutralizing solution, and centrifuged at $250 \times g$ for 10 minutes. The pellet was resuspended in 1 ml of KGM and PI (1 μ g) was added. After 3 minutes at room temperature, viability determinations were performed on the flow cytometer. Immediately after PI viability determinations, HEK DNA cell cycle analysis was conducted by adding PI-triton to each tube, and the samples were analyzed on the flow cytometer. Data analysis was performed as with the PBL.

SMITH, SANDERS, GALES, AND GROSS

NAD⁺ Determination:

Cellular NAD⁺ levels were determined spectrophotometrically by using the enzymatic cycling assay (11). Absorbance at 570 nm was measured by using a Spectronic 20 Spectrophotometer (Milton Roy, Rochester, NY) and compared to known amounts of NAD⁺. In all cases, the oxidized form of NAD (NAD⁺) was assayed. Previous work showed that the interconversion of NAD⁺ to NADH did not occur (7).

RESULTS:

HEK were more resistant to the cytotoxic effects of HD than were either quiescent or activated PBL (Figure 1). At low concentrations of HD (10 μ M or less), activated PBL were more sensitive than quiescent PBL. This difference was not seen at HD concentrations of ≥ 60 μ M. At the higher HD concentrations, there was a plateau attained with 30% survival for both resting and stimulated PBL.

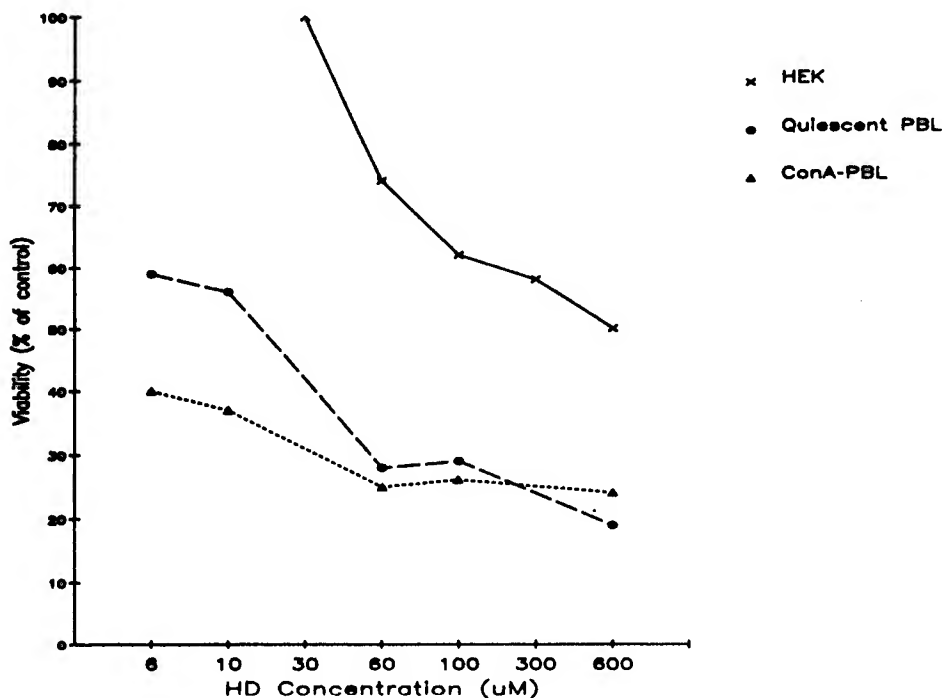


Figure 1. Viability response curves for HEK, quiescent PBL, and Con A stimulated PBL treated with various concentrations of HD. Viability was determined 24 hours after exposure by uptake of PI.

SMITH, SANDERS, GALES, AND GROSS

Niacinamide, an inhibitor of poly(ADP-ribose) polymerase, ameliorated the depletion of NAD⁺ in HEK following HD exposure (Figure 2). Control cells treated with 1 mM niacinamide demonstrated an increase of NAD⁺ to 110% of control within 4 hours of treatment. HD exposure resulted in a concentration-dependent reduction in cellular NAD⁺ to 30% of control at 600 μ M HD. At each concentration of HD tested, 1 mM niacinamide partially prevented this drop in NAD⁺ levels.

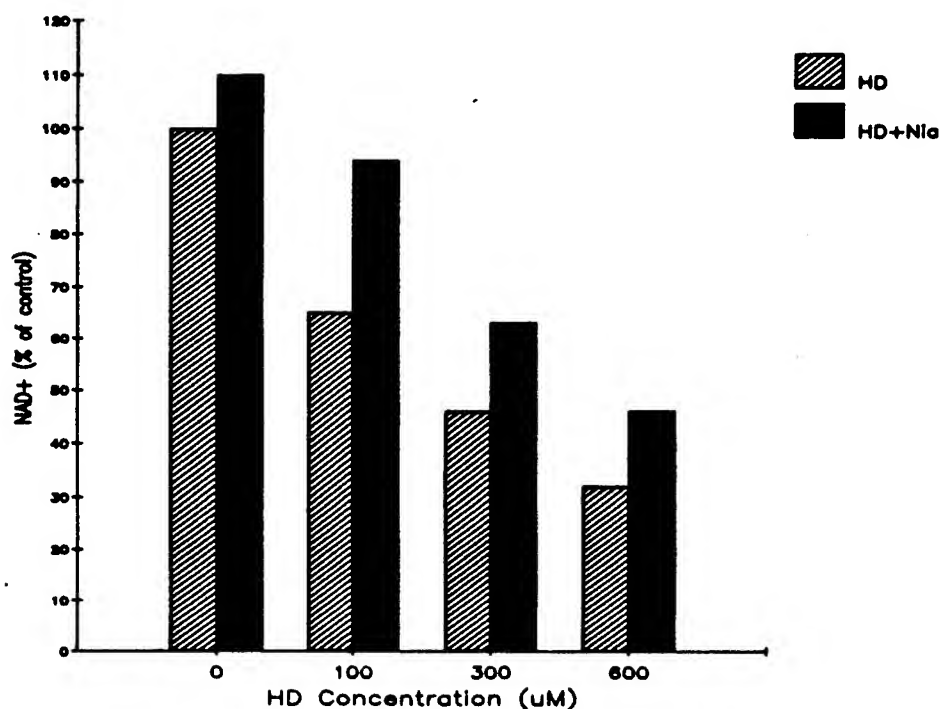


Figure 2. Ability of niacinamide (1 mM) to reverse the HD-induced depletion of cellular NAD⁺ levels in HEK. Cells were incubated in niacinamide at the time of HD addition and NAD⁺ levels determined 4 hours after HD.

Cytotoxicity from 300 μ M HD, when assessed by vital staining at 24 hours, could be prevented by niacinamide as shown in Table 1. Extended culture of these HD-treated cells in fresh media with or without 1 mM niacinamide resulted in loss of viability, as judged by lack of cell growth, in the HD-treated HEK within 72 hours.

SMITH, SANDERS, GALES, AND GROSS

Table 1. Survival of HEK (% of control) Following HD in the Presence or Absence of 1 mM Niacinamide

HD	Niacinamide	Viability at 24 hours	Extended Growth at 72 hours
---	---	100	100
---	+++	100	100
300 μ M	---	60	<5
300 μ M	+++	100	<5

Similar experiments with Con A-stimulated PBL showed protection by niacinamide up to 48 hours after HD exposure, but the protection decreased at 72 hours (Figure 3).

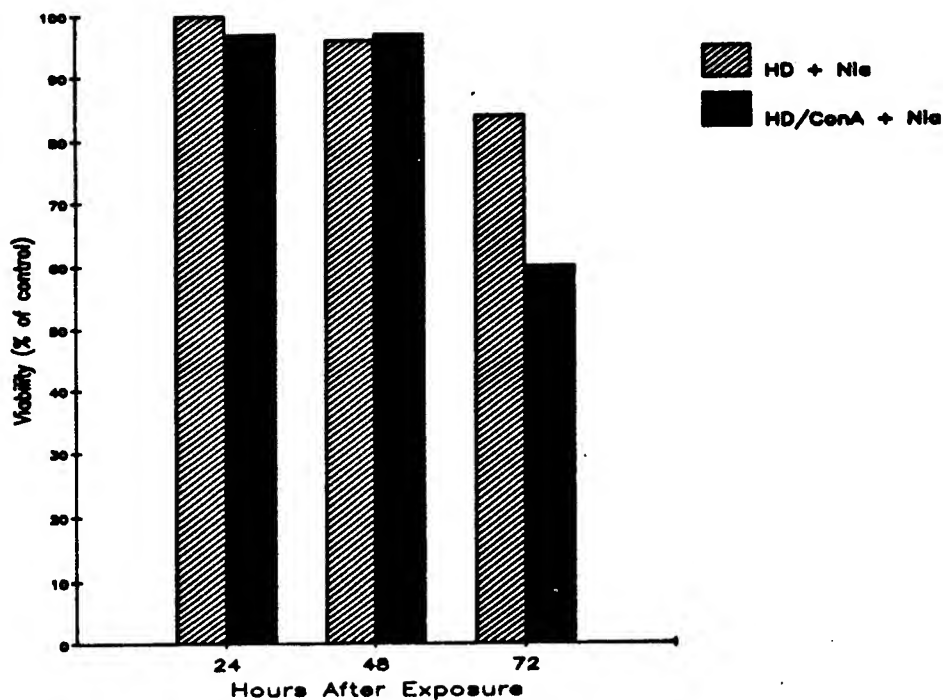


Figure 3. Ability of niacinamide (1 mM) to protect against HD-induced cytotoxicity in Con A-stimulated PBL. Cells were incubated in niacinamide at the time of HD (100 μ M) addition. Viabilities were determined at times shown after HD by uptake of PI.

SMITH, SANDERS, GALES, AND GROSS

When HD was added to activated PBL at various times after Con A addition, the increased cytotoxicity of cells in S phase could be seen (Figure 4).

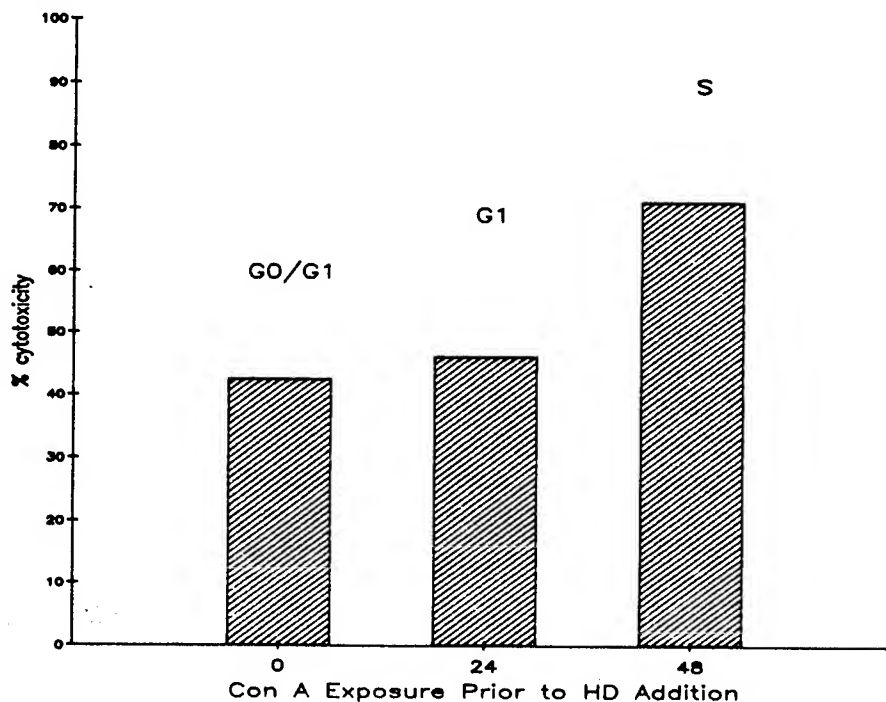


Figure 4. Cell cycle dependency of HD cytotoxicity on human PBL. Lymphocytes were stimulated by Con A for various periods of time prior to the addition of 300 μ M HD and tested 24 hours later for viability. Phase of the cell cycle for most of the cells at the time of HD addition is shown over the bars.

PI is a fluorescent dye which intercalates into the structure of DNA and binds stoichiometrically. The amount of dye taken up per cell increases as dividing cells progress through the replicative cell cycle. Flow cytometric analyses of dividing cells produce fluorescent histograms which, therefore, represent the full spectrum of DNA content seen in these populations. A typical DNA histogram for control HEK is shown in figure 5. These cells were evenly distributed throughout the cell cycle. DNA histograms of agent-exposed HEK indicated that the cell cycle was blocked in the late G1/early S phase (Figure 5).

The blockage of Con A stimulated PBL in early S phase could also be demonstrated by a blockage of ^3H -thymidine uptake 48 hours after activation. This blockage was not prevented by treatment with 1 mM niacinamide (Figure 6).

SMITH, SANDERS, GALES, AND GROSS

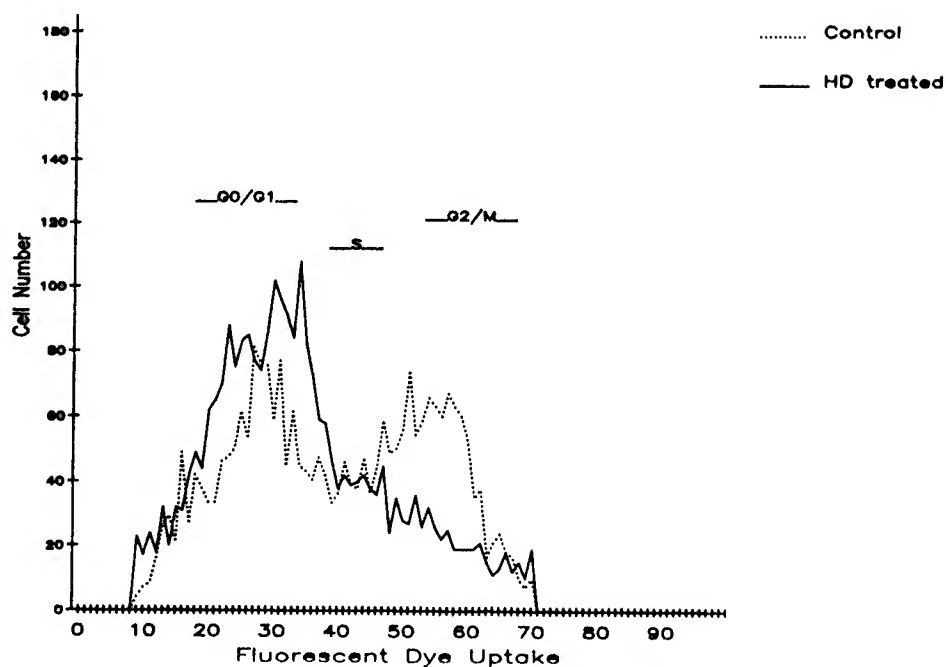


Figure 5. Blockage of HEK in S phase of the cell cycle by HD (600 μ M). DNA histograms were determined 24 hours after exposure by flow cytometric analysis of PI-triton staining as described in METHODS.

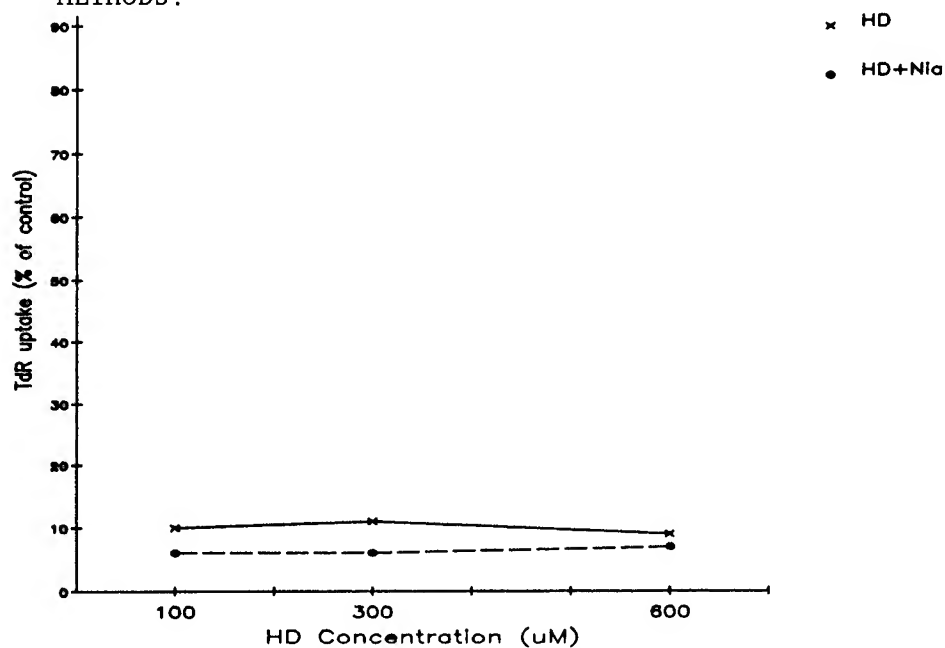


Figure 6. Failure of niacinamide (1 mM) to prevent the blockage of 3 H-thymidine (TdR) uptake in HD-treated, Con A-stimulated PBL.

SMITH, SANDERS, GALES, AND GROSS

In contrast to the fluorescent histograms generated from dividing cell populations, uptake of PI by the diploid DNA of resting PBL resulted in the detection of a single fluorescent peak. This peak was normally distributed and highly reproducible, varying less than 1 or 2 fluorescent channels in replicate samples. Within 4 hours following exposure to HD, and more markedly at 24 hours post-exposure, the treated cell DNA exhibited a distinct shift to the right, indicating increased dye uptake (Figure 7). By 48 hours post-HD, the DNA structure was totally destroyed as suggested by the complete inability (> 15 channel left shift) to take up PI.

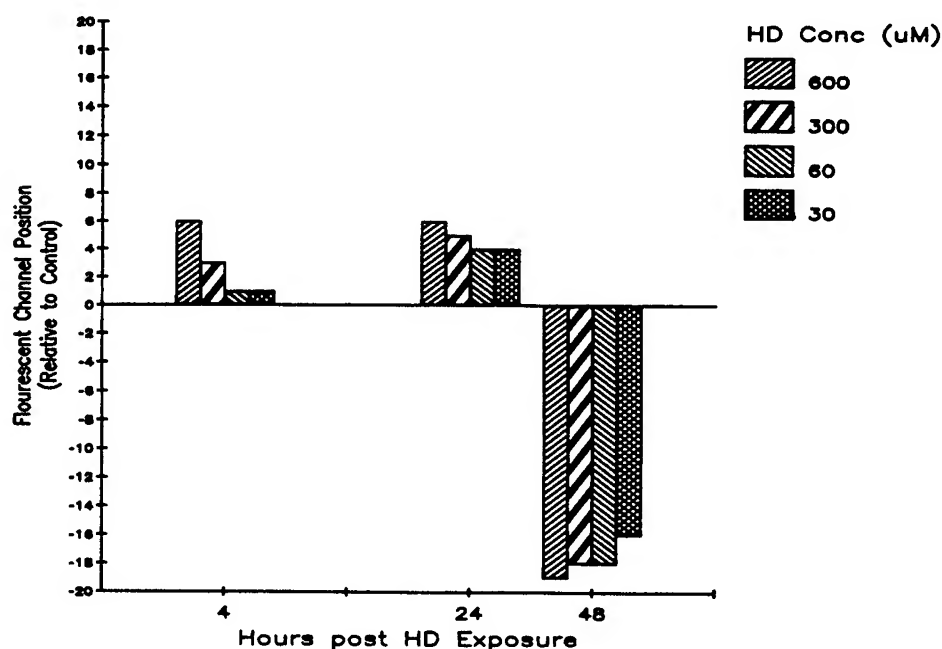


Figure 7. Shift of the peak diploid DNA fluorescent channel in HD-treated cells as compared to controls. Lymphocytes were treated with HD, and at indicated times after treatment, the cellular DNA was stained by PI-Triton and peak positions determined with Paral software. Positions positive to control indicate increased dye uptake. Positions negative to control indicate decreased dye uptake.

SMITH, SANDERS, GALES, AND GROSS

DISCUSSION:

All the cell types studied demonstrated concentration-dependent cytotoxic responses to HD, but HEK were less responsive than human lymphocytes. HEK are much larger than PBL with extensive cytoplasm and more glutathione than PBL. Since cellular DNA appears to be the major alkylation target of HD (5), this observation with HEK would suggest that they have more alternative alkylation sites than do PBL. This finding supports the results of Gross et al.(12) that suggests that increased levels of glutathione might protect cells against HD.

Activated PBL were more sensitive at low concentrations of HD, but this difference could not be seen at HD concentrations greater than 60 μ M. The increased sensitivity in dividing cells is not surprising since cells in S phase were the most sensitive to HD-cytotoxicity.

Niacinamide, which has been shown to protect human lymphocytes from NAD⁺ depletion and cytotoxic effects of HD (8), protected HEK against both effects for a short time. When the cytotoxic response was assessed by vital staining at 24 hours after HD, niacinamide not only protected treated cells but also preserved NAD⁺ levels, although not as completely as in resting lymphocytes (8). When the HEK cultures were permitted to grow in fresh media following HD exposure, the protective effect disappeared and the cells lost viability within 72 hours. The protective benefit of niacinamide did not completely wane in mitogen-stimulated PBL but was significantly diminished at 72 hours post-exposure. These data suggest that restriction of the HD-induced metabolic dysfunction is critical to the immediate post-exposure viability of cells. Further interventions, however, will be required to provide long-term survival following HD.

Proliferating cells allowed investigations into the effects of HD alkylation on cellular progression through the cell cycle. HD blocked cell cycling of both HEK and Con A-stimulated PBL in the early S phase. Along with the demonstration that cells in the S phase of the cell cycle were the most sensitive to the cytotoxic effects of HD, the data suggest that therapeutic approaches to HD should be directed towards maintaining cells in the G1 phase. This strategy would allow repair of DNA before the HD-exposed cells attempt DNA replication in the S phase.

Shifts in peak DNA fluorescent channel following HD exposure provide important information as to the repair capacity of alkylated DNA. Increased dye uptake at 24 hours post-exposure strongly suggests that the DNA is still undergoing repair processes. Since at 48 hours the lack of dye uptake suggests non-vital DNA, it can be assumed that failure to therapeutically interdict in the cellular DNA repair activities will be accompanied by loss of cell viability and subsequent pathology. On a more positive note, however, the DNA activity present at 4 and 24 hours after HD suggests an expansion of the therapeutic window for intervention well past the time formerly believed possible.

SMITH, SANDERS, GALES, AND GROSS

REFERENCES:

1. Cappizzi, RL, WJ Smith, RJ Field, and B Papirmeister. A Host-mediated Assay for Chemical Mutagens Using the L5178Y/Asn(-) Murine Leukemia. *Mutation Res* 21: 6, 1973.
2. Rozmiarek, J, RL Cappizzi, B Papirmeister, WH Furman, and WJ Smith. Mutagenic Activity in Somatic and Germ Cells Following Chronic Inhalation of Sulfur Mustard. *Mutation Res* 21: 13, 1973.
3. Smith, WJ, KM Sanders, CL Gross, and HL Meier. Flow Cytometric Assessment of the Effects of Alkylation on Viability, Cell Cycle, and DNA Structure in Human Lymphocytes. *Cytometry Suppl* 2: 21, 1988.
4. Vedders, EB. The Vesicants, in The Medical Aspects of Chemical Warfare. Baltimore, MD, Williams and Wilkins Co. 1925, pp 125-166.
5. Papirmeister, B, CL Gross, HL Meier, JP Petralli, and JB Johnson. Molecular basis for mustard-induced vesication. *Fund and Appl Toxicol* 5: S134-S149, 1985.
6. Gross, CL, HL Meier, B Papirmeister, FB Brinkley, and JB Johnson. Sulfur mustard lowers nicotinamide adenine dinucleotide concentrations in human skin grafted to athymic nude mice. *Toxic Appl Pharmacol* 81: 85-90, 1985.
7. Meier HL, CL Gross, and B Papirmeister. 2,2'-Dichlorodiethyl sulfide (sulfur mustard) decreases NAD+ levels in human leukocytes. *Toxicol Lett* 39: 109-122, 1987.
8. Meier, HL, CL Gross, LM Graham, CT Lusco, and JB Johnson. The prevention of 2,2'-dichlorodiethyl sulfide (sulfur mustard, HD) cytotoxicity in human lymphocytes by inhibitors of poly(ADP-ribose) polymerase. *Proc of the 6th Med Chem Defense Bioscience Rev*: 313-316, 1987.
9. Meier, HL, JP Petralli, and CL Gross. Niacinamide prevents sulfur mustard (HD)-induced pathology and biochemical changes in human lymphocyte and mixed human leukocyte preparations. *Proceedings of the 1988 Army Science Conference*.
10. Taylor, IW. A rapid single step staining technique for DNA analysis by flow microfluorimetry. *J Histochem Cytochem* 28: 1021-1024, 1980.

SMITH, SANDERS, GALES, AND GROSS

11. Jacobson EL and MK Jacobson. Pyridine nucleotide levels as a function of growth in normal and transformed 3T3 cells. Arch Biochem Biophys 175: 627-634, 1976.
12. Gross, CL, JK Innace, RC Krebs, WJ Smith, and HL Meier. Sulfur mustard (HD) cytotoxicity in lymphocytes can be affected by intracellular glutathione levels. FASEB J., 3: 583, 1989.

Motion Planning for the
Universal Self-Deployable Cargo Handler (USDCH) (U)

* Stephen F. Sousk, Dr.

U.S. Army Belvoir Research, Development and Engineering Center,
Fort Belvoir, VA 22060-5606

Francois G. Pin, Dr.

Autonomous Robotic Systems Group

Oak Ridge National Laboratory

Oak Ridge, TN 37831-6364

Carl D. Crane, Dr.

Center for Intelligent Machines and Robotics

University of Florida

Gainesville, Florida 32611

INTRODUCTION

Belvoir Research, Development and Engineering Center (Belvoir) is developing advanced material handling equipment for sustaining materiel requirements of AirLand Battle-Future. In particular, Belvoir is pursuing the development of a rough terrain forklift truck, the Universal Self-Deployable Cargo Handler (USDCH), that also is called ATLAS.

The USDCH is a rough terrain materials handling vehicle that will perform support functions on a continuous basis in hazardous/hostile environments and will, in its first phase, incorporate convoy mobility. It is intended as a replacement for the Army's current family of 4000, 6000, and 10000 pound lifting capacity rough terrain forklift trucks. The hardware incorporates features to allow economic preplanned product improvement (P3I) that will lead to fully autonomous operations by the year 2015.

The approach of the USDCH development program is to allow interaction between the operator and the automated controlling systems, while progressively adding robotics, artificial intelligence, and automated methods to achieve a high level of autonomy. Major challenges involve the development of methodologies to provide real-time operations in hostile field environments and to exploit human-machine synergism to obtain maximum efficiency and reliability.

The USDCH is a car-like wheeled platform supporting a manipulator that has redundant degrees of freedom in its plane of manipulation. This paper reports on robotic motion planning for two major functions of the USDCH: platform mobility and motion of the material handling arm.

Novel approaches were necessary to automate the basic motions of the vehicle and the arm. The algorithms for positioning the platform at a work area are presented and results showing optimum vehicle trajectories are given. These results show that the methodologies and algorithms developed are applicable to a wide class of vehicles. The kinematics governing the planning of the redundant arm motion are presented, and the strategies for controlling the arm in various cargo acquisition tasks are discussed.

MOTION PLANNING FOR PALLET ACQUISITION

The automation of the USDCH will occur in incremental stages. The first stage is robotically-assisted pallet acquisition. In an operational scenario, the operator drives the USDCH "close" to a pallet, "designates" the pallet, and the pallet locating sensors identify the pallet position and orientation relative to the USDCH vehicle base. With this information, motion planning algorithms plan the required movements of the vehicle and the arm to acquire the pallet and move it to a travel position.

At the pallet drop-off point, a similar process occurs. The operator will be able to specify a downloading pattern, drive "close" to the drop-off point, and then initiate the autonomous downloading sequence, to include motions of the platform to reach an appropriate location and movement of the arm to lay down the load. Automation of the USDCH platform motion is thus necessary in all material handling and transfer tasks, and automation of the USDCH arm motion is a common element of the pallet acquisition and downloading problem. An analysis of the kinematic nature of the USDCH, as well as the motion planning strategies to be used for platform and arm motion in pallet acquisition tasks is discussed in the following sections of this paper.

IMPLICATION OF THE USDCH KINEMATICS ON PALLET ACQUISITION STRATEGY

An obvious first step in the USDCH motion analysis is to obtain a closed form reverse kinematic analysis for the arm. The USDCH arm contains several degrees of freedom as illustrated in the kinematic diagram shown in Fig. 1. The figure depicts the USDCH with the 6000 or 10000 pound lifting capacity forks attached. In either of these configurations, the arm contains eight independent freedoms, i.e., rotations about the axes S_1 , S_3 , S_5 , S_6 , S_7 , and S_8 and translations along the vectors S_2 and S_4 . The two fork tines are each independently controlled and can be translated along the S_9 vector. Thus the USDCH arm, with either the 6K or 10K fork attached, has ten independent freedoms. A third fork attachment which can lift 4K loads includes a scissor mechanism that allows translation as well as rotation along the S_8 vector. The USDCH arm, configured with the 4K fork attachment, thereby possesses eleven freedoms.

A closer analysis of the USDCH kinematics reveals that in most situations, a desired position and orientation of the fork tines is outside the reachable workspace of the arm. This is due in large part, to the small angle of rotation about the first joint axis ($\pm 7^\circ$ from the position shown in Fig. 1). For most applications, rotation about S_1 is undesirable, and this joint will be locked.

It also can be seen that joints S_2 through S_6 perform planar motion. With the joint 1 locked, these joints and the wrist can only move in one plane of the vehicle coordinate system. These five joints are redundant because only three freedoms are required for planar motion.

An analysis of the nature of the pallet acquisition problem places further restrictions on the system. The elasticity of the tires may cause large, undesirable motions when rolling on rough terrain during the pallet engagement procedure. It is desirable that the vehicle base not move while the fork accomplishes load acquisition or lay down. The vehicle first must be oriented so the vector S_8 lies in the X-Z plane of Fig. 2 and the seventh joint is locked at zero degrees. With these kinematic and operational constraints, the USDCH arm effectively behaves like a redundant planar manipulator; therefore, it is critical that the vehicle base be positioned and oriented properly so that the pallet can be acquired. The determination of this position and orientation and the planning of the vehicle motion needed to move the vehicle base to that configuration represent essential components of robotic pallet acquisition by the USDCH.

DETERMINATION OF VEHICLE POSITION AND ORIENTATION FOR PALLET ACQUISITION

It is assumed that the sensors provide the position and orientation of a pallet in the world coordinate system. This data can readily be transformed so that the necessary position of the point O_6 and the orientation vector S_8 in Fig. 1 for an approach position (before engagement) and a pick-up position (fork fully engaged in pallet) are known.

The USDCH platform must now be oriented to put the vector S_8 in the X-Z plane of the vehicle coordinate system. This requirement is necessary so that the forks can be moved from the approach point to the pick-up point without moving the vehicle base itself. Shown in Fig. 2 is a drawing of the USDCH forks displayed with the X-Z plane of the vehicle coordinate system. The orientation angle of the X-Z vehicle plane, γ , can be determined from simple projection of S_8 onto the X-Y plane:

$$S_{8xy} = S_8 \circ \underline{i} + S_8 \circ \underline{j}$$

Once this projected vector is unitized, its x component will equal the cosine of γ . This projection technique will be indeterminate if the vector

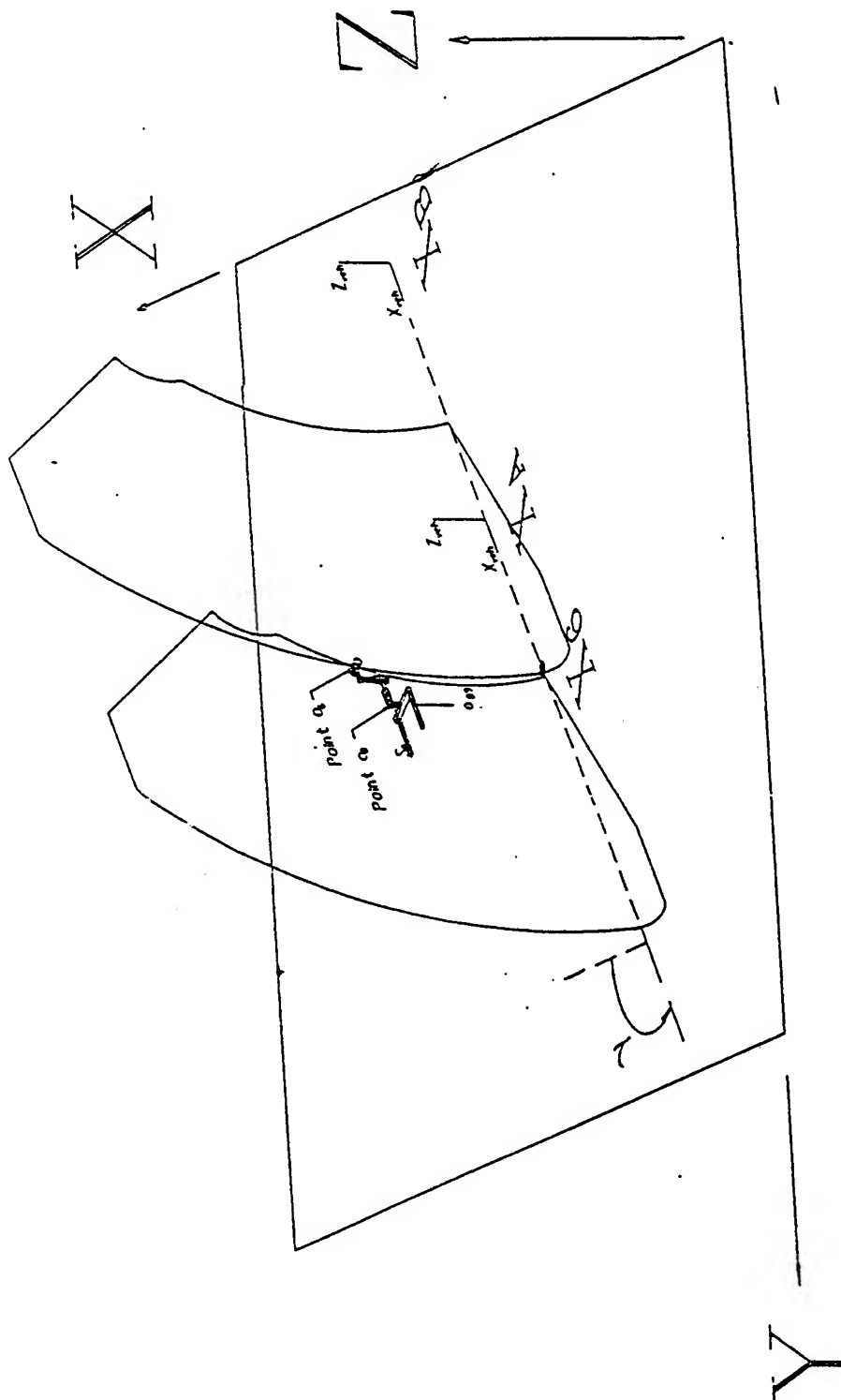


Fig. 2. Determination of allowable region for USDCH base coordinate system.

S_8 is parallel with the Z axis, which is not a possibility for the pallet acquisition problem.

The exact location of the X-Z plane of the vehicle coordinate system as measured in the world coordinate system is known since the orientation of the plane, γ and the location of one point on the plane (e.g., point O_6) are known. In Fig. 2, point O_6 has been projected onto the world X-Y plane and is labeled as point X_6 . The question remains: how far from point O_6 can the origin of the vehicle coordinate system lie? This can be answered by investigating the effective workspace of point O_6 with respect to the vehicle coordinate system. This planar workspace is determined by moving each of the rotational joints (S_3 , S_5 , and S_6) and each of the translational freedoms (S_2 and S_4) from their minimum to their maximum values. Fig. 2 shows workspace comprised of eight straight line or curved segments. The analytic equations of these segments are derived from the vehicle coordinate system, based on the kinematic models of the USDCH.

With the workspace of point O_6 known, it is now possible to determine the allowable range of locations of the vehicle base so that the arm can reach the pallet. Figure 2 shows the workspace of point O_6 in its two extreme positions. A computer subroutine calculates the distance between the projection of point O_6 on the X axis of the vehicle coordinate system, i.e., point X_6 and the extreme locations for the origin of the vehicle coordinate system. This distance data, together with the coordinates of point O_6 measured in the world coordinate system and the orientation angle, γ , can be used to calculate the extreme allowable positions for the vehicle base in terms of the world coordinate system. The output of the algorithm is the orientation angle γ and the coordinates of points X_A and X_B as shown in Fig. 2.

These results now can be transmitted as input to the vehicle motion planning algorithms (discussed in the next section) to calculate the precise path of the vehicle to the position required for pallet acquisition.

PLATFORM MOTION PLANNING

The USDCH vehicle base provides excellent mobility via a tight turning radius and a "crab steer" capability. As a mobile system moving in a 2-D world, it has 3 robotic degrees-of-freedom: the coordinates of a given point of the robot, and the orientation of the robot itself. When path planning is part of a task to be performed by the robot, the 3 degrees of freedom must be taken into account. Therefore, a path is a curve joining 2 points in space and providing a prescribed change in the robot orientation.

Unfortunately, a path depends on the motion characteristics of the robot. Car-like robots, such as the USDCH, cannot follow an arbitrary path because of the non-holonomic constraints imposed by the kinematics of

their platform. The method described below provides a path joining any couple of configurations for a car-like robot.

Let us denote by L the distance between the 2 axles of the wheels; by θ the angle between the major axis of the robot and the x axle; by ϕ the steering angle (i.e., the orientation of the front wheels with respect to the major axis of the robot); and by x_M , y_M and V_M the coordinates and the velocity of M , the middle of the axle of the rear wheels.

We assume that there is no slipping of the rear wheels, that the angle of the velocity of M with respect of the x axle is θ . Hence:

$$\begin{pmatrix} \frac{dx_M}{dt} \\ \frac{dy_M}{dt} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

are collinear. Therefore:

$$\frac{dx_M}{dt} * \sin \theta - \frac{dy_M}{dt} * \cos(\theta) = 0$$

This equation is not integrable. It is a constraint on the velocity of the robot, but it does not affect the dimension of the space of the configurations (the vehicle can have any orientation at any position). On the other hand, for a given configuration, the dimension of the space of achievable velocities is only $2 - 1 + 1$ (the velocity must have the same orientation as the robot). This is a non-holonomic constraint.

Another characteristic of the USDCR platform is that the steering angle is limited:

$$|\phi| \leq \phi_{\max} < \frac{\pi}{2}$$

Combining the velocity equations of the vehicle with the non-holonomic constraint, we can obtain:

$$\frac{ds}{d\theta} = \frac{L}{\sin \phi}$$

where s is the curvilinear distance along the path. In terms of R , the radius of curvature of the trajectory:

$$|R| \geq R_{\min} = \frac{L}{\sin \phi_{\max}}$$

PRINCIPLE OF THE ALGORITHM

The problem of deterministically finding trajectories for mobile robots with non-holonomic constraints such as the USDC has never been formally solved. Barraquand and Latombe used a discretized configuration space domain to find some trajectories.¹ Their method, however, is computationally expensive and, due to the discretisation, provides trajectories that are accurate only to within a mesh size. Loss, Boulle and Tournassoud proposed a set of trajectories composed of two arcs of tangent circles.² It is a good method for medium range goals, but for close goals (e.g., positioning for pallet acquisition) or far goals, the method is not applicable since the trajectories are sometimes prohibitively long. Kanayama and Miyake³ have used clothoids (curves with linear variations of the curvature) to design trajectories, the drawback of which is the complicated geometry. When obstacle avoidance is involved, the calculation of the distance between the planned trajectory and a potential obstacle is required and is computationally complex with clothoids. The method which is proposed here solves the trajectory planning problem deterministically, with a limited amount of needed computations, making it particularly useful for real-time sensor-based applications.

The constraint imposing a radius of curvature greater than R_{\min} for the trajectory restricts the shape of the extremities of the trajectory: to reach the configuration drawn in Fig. 3, "the end" of the trajectory must remain outside the two circles on the sides of the final position of the robot.

More precisely, two such circles can be associated to both the initial configuration and the final configuration. If the robot enters one of the circles related to the final configuration, it first will have to get out of it before being able to reach the final configuration. Similarly, from the initial configuration, in order to enter one of the circles, the robot must go away from them first.

If it provided the desired change in orientations, the best path between two configurations would be a straight line segment. On the other hand, the fastest way to change the orientation of the robot is to move along a circle of radius R_{\min} . Therefore, the trajectories are designed in the following way: the robot moves along one of the circles related to the initial configuration in order to be oriented towards the goal, no matter what the final orientation must be. Then, it moves along a straight line segment in the direction of the goal. The last part of the trajectory is meant to provide the desired final orientation: the robot moves along one of the circles related to the final configuration. To be physically achievable, the trajectory must provide a continuous

orientation to the robot. Therefore, the straight line segment must be tangent to the circles (see Fig. 4).

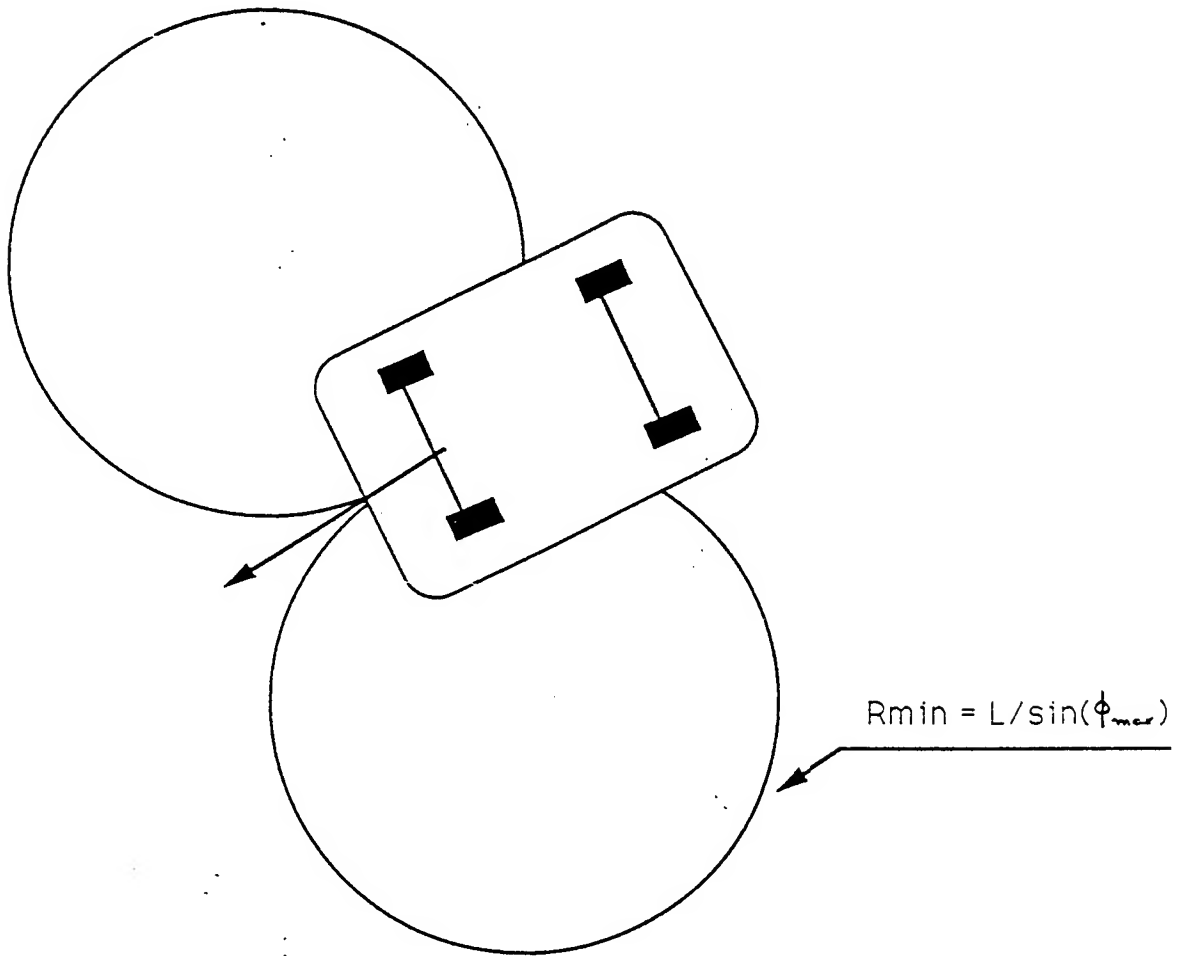


Fig. 3. Circles corresponding to a minimum radius of turn for a given configuration.

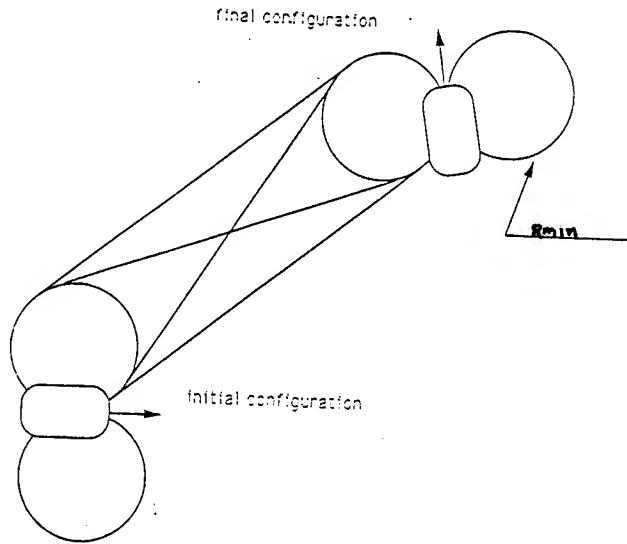


Fig. 4. Possible trajectories between two configurations.

If the distance between the center of two circles is greater than the sum of their 2 radii, there are 4 common tangents to the circles. Since there are 4 possible couples of circles, there are 16 possible paths. Only 8 paths provide the desired orientation, the others lead to the opposite orientation. Among the 8 remaining paths, the shortest one is chosen; and it can be shown that this trajectory is optimal for minimizing travel distance (see Fig. 5).

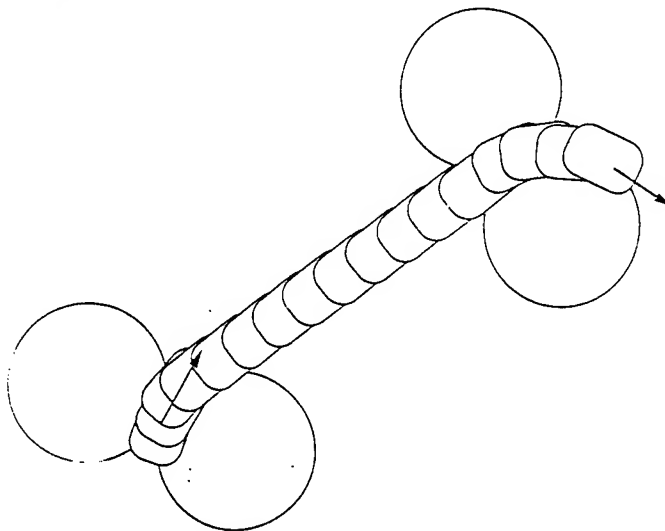


Fig. 5. An optimum trajectory between two configurations.

If the distance between the 2 chosen circles is smaller than $2 \cdot R_{\min}$, there are only 2 common tangents, but the method remains the same.

When the 2 positions of the robot are very near each other, backing up and maneuvering are often involved. In this case, the method also provides the maneuvers in a straight forward manner (see Fig. 6).

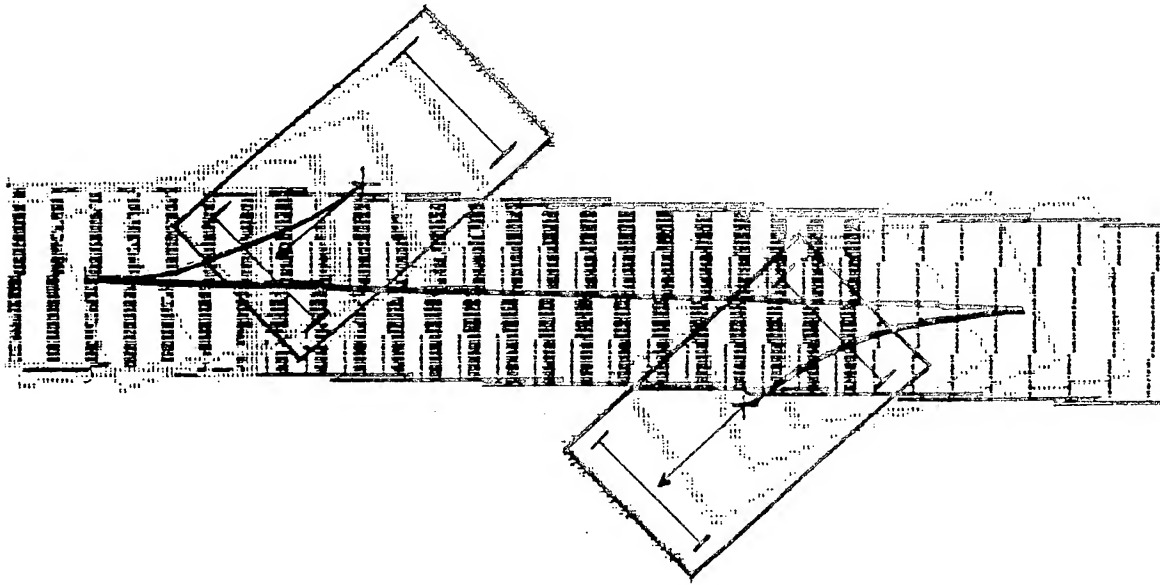


Fig. 6. A trajectory involving forward and reverse maneuvering.

This algorithm provides a continuous orientation for the robot because the circles and the segment are tangent. However, the curvature along the curve is discontinuous: it is a non-zero constant on the circles and zero on the straight line segment. From a mechanical point of view, it means that the orientation of the wheels changes suddenly at one point. When the robot leaves the circle, the steering angle is ϕ_{\max} , on the straight line, it is null. This does not mean that the trajectory is not achievable. It only means that since the robot cannot turn its wheels instantaneously, it has to stop and resume moving when the steering angle is correct. This is already the case when the robot has to maneuver and reverse its velocity, which is a common occurrence in an obstructed environment. For the other cases, we now describe a method to provide continuous curvatures along the trajectories.

TAKING INTO ACCOUNT THE STEERING RATE

In this section, the hypotheses about the robot are the same as in the previous section except that an additional constraint concerning the steering angle is imposed on the system:

$$|\dot{\phi}| \leq \dot{\phi}_{\max}$$

Let us assume that the vehicle is moving at a velocity equal to V and that the steering angle of the front wheels is ϕ_{\max} . Let us also assume that the steering angle and the velocity of the robot are varying as:

$$v = V + \gamma * t$$

$$\phi = \phi_{\max} - \dot{\phi}_{\max} t$$

Then, when the robot changes its steering angle from ϕ_{\max} to 0, the equations of the trajectory are:

$$x_N(s) = \int_0^s \cos(\theta) ds$$

$$y_N(s) = \int_0^s \sin(\theta) ds$$

The interesting point is that since the robot moves at a constant velocity on the circle, the relative trajectory at any point on the circle where the robot initiates the change in steering angle from ϕ_{\max} to 0, is given by the above equations and can be computed off line and stored in the robot memory.

Now, in order to find a path requiring a continuous steering angle, we have to find the lines intersecting 2 circles with a given angle (always the same). The resulting trajectories (see Fig. 7) then satisfy the non-holonomic kinematic constraint, the maximum steering angle constraint, and a bounded rate of change of the steering angle (the latter characteristic allowing compatibility with a bounded control speed for autonomous steering of the vehicle).

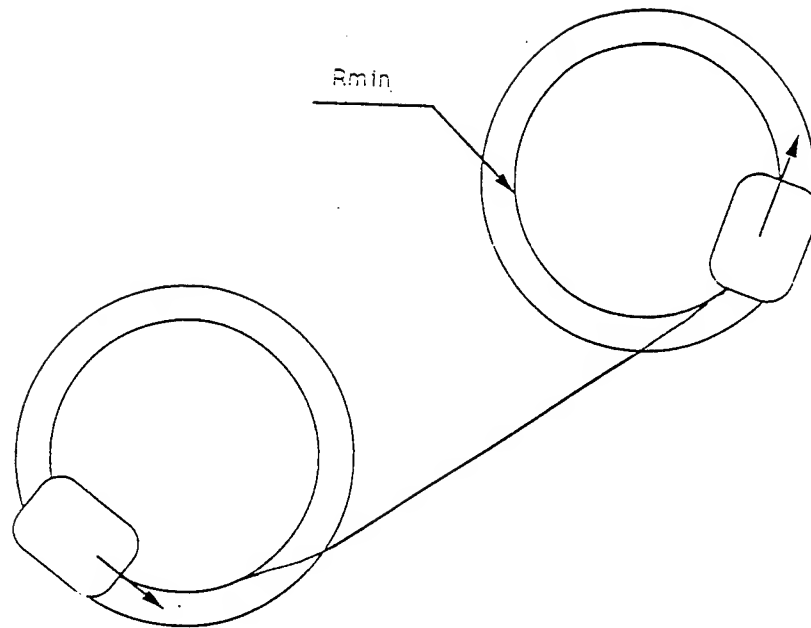


Fig. 7. Trajectory between two configurations with non-holonomic constraint, maximum steering angle constraint and a bounded rate of change of the steering angle.

ADVANTAGES OF THIS METHOD

The main advantage of the method is that it provides a trajectory joining all pairs of configurations of the robot in space. The method also allows plan trajectories requiring maneuvering, taking into account the forward and reverse mode of motion of the vehicle or a combination of both. The algorithm can be easily used by an operator who just specifies the desired final configuration of the vehicle. When the robot is used in an autonomous mode, the method can be used as a basic motion planner even in obstructed environments. The fact that the trajectories all have the same composition, with well known geometric elements, enables easy and fast computation of the distances in relation to the obstacles. Therefore, this method provides a good base for real-time local and global navigation of the USDCH in unstructured environments.

CONCLUSIONS

This paper analyzes robotic motion planning for two major functions of the USDCH: platform mobility and motion of the material handling arm.

Strategies for the redundant arm motion for pallet acquisition tasks have been presented. The problem of deterministically computing trajectories for mobile robots with two-axle, car-like, non-holonomic constraints has been solved.

Previous attempts at formally solving this problem have been unsuccessful. The solutions needed for mobile manipulators such as the USDCH have never been obtained before.

The methods described in this paper are readily extendable to the more general problem of real-time sensor-based planning in unknown and unstructured environments. The methods are applicable to a wide range of material handling vehicles. For the USDCH, the methods represent an important step toward the goal of robotic-assisted material handling using a rough terrain forklift truck. Additional applications exist for other military systems and for hazardous waste management.

ACKNOWLEDGEMENT

The first author acknowledges the contributions to the USDCH program of David Krawchuk, Trung Van, and Donna Dodge, Logistics Equipment Directorate, BELVOIR.

REFERENCES

1. Barraquand, Latombe, "On Nonholonomic Mobile Robots and Optimal Maneuvering," 4th IEEE International Symposium on Intelligent Control, Albany, NY, 1989.
2. Loss, Boulle, Tournassoud, "Local Trajectory Generation for a Mobile Robot," INRIA, France (internal paper).
3. Kanayama, Miyake, "Trajectory Generation for Mobile Robots," Proceedings 3rd International Symposium of Robotics Research, Gouvieux, France, 1985.

SPARENBORG, BRENNECKE, & BRAITMAN

The Excitatory Amino Acid Antagonist MK-801 Prevents
Nerve Agent-induced Neuropathology Even When Given
After the Onset of Convulsions

Steven Sparenborg, CPT, Lucas H. Brennecke, MAJ
and David J. Braitman, MAJ
U.S. Army Medical Research Institute of Chemical Defense
Aberdeen Proving Ground, MD 21010-5425

At the close of World War II, the allied powers learned that Germany had developed a frighteningly potent chemical weapon which they called G-gas. This substance is an organophosphorus agent whose chief mechanism of action is the irreversible binding of acetylcholinesterase, an enzyme responsible for metabolizing the neurotransmitter acetylcholine (ACh). Lethal doses of organophosphorus nerve agents produce death by respiratory arrest, but severe convulsions occur prior to death. With the use of today's medical countermeasures of pyridostigmine, atropine, and pralidoxime chloride to minimize the fatal outcome of nerve agent poisoning, controlling convulsions is now of chief concern.

Overlooked for more than three decades by those studying nerve agents, brain damage was reported first by Petras (1981) in rats which survived for many days after severe soman poisoning. Neuronal damage was observed in many parts of the brain. Subsequent investigations have described a consistent pattern of severe neuronal necrosis, chiefly in the hippocampus, amygdala, pyriform cortex, thalamus, and cerebral cortex (Lemerrier et al., 1983; McLeod et al., 1984). Neuronal necrosis is the endpoint of a pathological process requiring many hours to complete, and is found most often in animals exhibiting severe convulsions. Therefore, the possibility of brain damage resulting from soman-induced convulsions increases the need for an anticonvulsant which will completely protect against convulsions and brain damage.

The first effective anticonvulsant for nerve agents was identified by Lipp (1972). This Canadian defense researcher found that diazepam (Valium), the standard drug for arresting status epilepticus in humans, was also effective in preventing and arresting the convulsions caused by the nerve agent soman (GD) in monkeys when administered at the first sign of brain seizure activity as measured by EEG. Martin et al. (1985) were the first to demonstrate that diazepam would prevent brain damage in rats poisoned with soman. These researchers injected diazepam 30 minutes before the injection of soman. Because diazepam has long been known to prevent soman-induced convulsions and brain damage, when given before

SPARENBERG, BRENNECKE, & BRAITMAN

convulsions have fully developed, it has been at the forefront of compounds considered by military forces throughout the world for use as an anticonvulsant for nerve agents.

The tranquilizing effects of diazepam would produce unacceptable decrements of performance if given as a pretreatment in advance of a chemical attack. Its use as an anticonvulsant for nerve agents must, therefore, be restricted to post-attack administration to individuals exhibiting convulsive symptoms. Little is known about the efficacy of diazepam when its administration is delayed until after convulsions have developed. We have designed this study to determine, if possible, the critical period after which diazepam injection would no longer protect against convulsions and brain damage.

In addition we have attempted to identify novel approaches to protect against nerve agent-induced convulsions and brain damage. Increasing evidence indicates that endogenous excitatory amino acid (EAA) neurotransmitters, such as glutamate and aspartate, play a key role in the initiation and propagation of seizure activity and in brain damage induced by ischemia, anoxia, and seizures. In particular, the N-methyl-D-aspartate (NMDA) subtype of excitatory amino acid receptors has been identified as a site of action for these pathological phenomena. Consistent with these observations, NMDA antagonists are potent anticonvulsants (Clineschmidt et al., 1982; Gilbert, 1988; McNamara et al., 1988) and protect against brain damage induced by several insults (Foster et al., 1987; Kochhar et al., 1987; Westerberg et al., 1988).

Braitman and co-workers (Braitman et al., 1988, 1989) demonstrated that the experimental drug MK-801 ((+)-5-methyl-10, 11-dihydro-5H-dibenzo [a,d]cycloheptan-5,10-imine maleate), one of the most potent and selective antagonists of the NMDA class of receptors (Wong et al., 1986), protected guinea pigs against soman poisoning. MK-801, given 30 minutes before soman challenge, was found to reduce the duration and intensity of seizures and convulsions, prevent neuronal necrosis, and reduce the incidence of death resulting from highly toxic doses of soman.

Although MK-801 gave excellent protection in those experiments, its efficacy when given after the onset of convulsions had not been tested. By giving MK-801 or diazepam to guinea pigs experiencing soman-induced convulsions at varying times after convulsion onset, one might learn when the process of necrosis became irreversible by pharmacological intervention. Such information may be useful in the selection of an anticonvulsant for military use.

Methods

Subjects and surgical preparation

Male albino Crl:(HA)BR VAF/Plus guinea pigs (*Cavia porcellus*) weighing 300-350 g were quarantined on arrival and screened for evidence of disease. They were individually housed in polycarbonate shoebox cages on

SPARENBORG, BRENNECKE, & BRAITMAN

contact corn cob bedding changed two times per week and provided with commercial certified guinea pig ration (Purina # 5026, Purina Mills, Inc. Richmond IN 47375) and tap water ad libitum. Animal rooms were maintained at 20-22°C, relative humidity of $50 \pm 10\%$, on a 12-hour light:dark cycle with no twilight. Each subject was chronically instrumented for recording of the electrocorticogram (ECoG) using standard methods as described in an earlier report (Braitman and Sparenborg, 1989).

Experimental design and drugs

The guinea pigs were placed in individual, plexiglas cubicles for ECoG recording and experimental drug administration. The experiment began with a 30-minute baseline period of behavioral observation and ECoG sampling, followed by a 30 minute pretreatment period. At the beginning of the pretreatment period, animals received an i.m. injection of pyridostigmine (0.13 mg/kg). The pretreatment period was followed immediately by the soman treatment period. At the beginning of the soman treatment period, each subject received 2xLD₅₀ soman (56 ug/kg, s.c.). Pralidoxime chloride (25 mg/kg), more commonly known as 2-PAM, and atropine sulfate (4 mg/kg) were given (i.m.) therapeutically at 15 sec and 2 min, respectively, after soman. The 2xLD₅₀ dose of soman was used to produce unequivocal seizures in each saline-treated subject. Drug doses were selected to increase the probability of robust convulsions by providing maximal protection against peripheral respiratory arrest, but only minimal protection against central seizures.

A test drug was administered (i.m.) to guinea pigs 5, 50, or 120 minutes after soman-induced spikes first appeared in the oscilloscope record. MK-801 (0.3 mg/kg) or diazepam (1 mg/kg) was given at all three injection delay times. Higher doses of the drugs were given at the 5- and 120-min injection delay times (MK-801, 1 mg/kg; diazepam, 3 mg/kg). Two control groups were injected with the test drug vehicles (MK-801 = normal saline; diazepam = ethanol, propylene glycol, and benzyl alcohol in water) at the 5-min injection delay time and are referred to as the diazepam- or MK-801-control groups. Two additional control groups were given pyridostigmine, 2-PAM, atropine and one of the test drugs at the higher dose, but in place of soman, an injection of saline was given. In these two soman-control groups, the test drug was given at 15 min after the saline injection. All subjects were observed for 6 hours immediately following the soman injection time. Each treatment group consisted of 8-10 animals.

Data collection and analysis

During testing, the difference voltage from the two electrodes was amplified and band-pass filtered (half amplitude cut-offs: 0.1 and 100 Hz). The signal was then passed to an analog-to-digital converter, an FM tape recorder, an oscilloscope, and a paper chart recorder. The paper chart records were examined for electrographic seizure activity (EGSA),

SPARENBERG, BRENNECKE, & BRAITMAN

characterized by repetitive spikes or sharp waves (2-5/sec). The digitized signal was sampled 4 times per minute (sample epoch = 2 sec) and a fast Fourier transform performed on each epoch. The transformed signals were then called power spectral estimates (PSE) which reflect the amplitude and frequency characteristics of the ECoG signal at given time points.

The magnitude of ECoG activity changes between conditions was standardized for individual subjects by converting the PSE to z-scores. To accomplish this, the means of the baseline period PSE were subtracted from the pretreatment and soman period PSE and then divided by the baseline period PSE standard deviation.

PSE were analyzed for three periods after soman administration. The first period began 5 minutes before, and ended simultaneously with, the injection of the test drug. The second period began with the injection of the test drug and was followed immediately by the third period. Each were 15 min in duration.

Drug treatment group z-scores for spectral power were compared by two-way ANOVA's. Paired comparisons of treatment group means were made with the Fisher Least Significant Difference test following significant outcomes in the ANOVAs.

Neuropathology

Those guinea pigs surviving at 48 hours after the administration of soman were deeply anesthetized with pentobarbital sodium and transcardially perfused with saline followed by 10% formalin. The brains were then removed and processed for hematoxylin and eosin staining.

Soman-induced brain lesions are known to occur primarily in the pyriform cortex, hippocampus, thalamus, cerebral cortex, and amygdaloid nuclei (McLeod et al., 1984). Three standard brain sections that permitted concurrent evaluation of these sites were chosen and collected for histopathologic evaluation. A frontal tissue block section was taken through the center of the infundibulum and 2 mm caudal to the infundibulum. Sections (thickness = 6 microns) were placed in embedding cassettes with the rostral side down so that identical sections were evaluated in all test subjects.

The sections from each subject were evaluated by a certified (DACVP) veterinary pathologist (L.H.B.) who was unaware of the group assignment of individual subjects. A numerical score was given for the severity of neuronal necrosis in each structure using the following key: 0 = no visible neuronal necrosis, 1 = minimal change with < 5% neuronal necrosis, 2 = mild change with 6-15% neuronal necrosis, 3 = moderate change with 16-30% neuronal necrosis, 4 = severe with > 30% neuronal necrosis. The distribution of damage was also scored with a five-point system (0 = no lesion, 1 = focal, 2 = focal but extensive, 3 = multifocal, 4 = diffuse). The product of severity and distribution necrosis scores for each brain structure in each subject were computed.

SPARENBORG, BRENNER, & BRAITMAN

Thus, the maximum score of 16 represented the most necrotic tissue. The Kruskal-Wallis test was used to perform analyses for each brain area comparing the no-soman group to the three groups receiving soman (5, 50, and 120 min injection delay times) for each test drug. Selected paired comparisons of groups were performed, using the method described by DeVore (1982), following significant outcomes of the Kruskal-Wallis test. All statistically significant differences reported here are based on outcomes with $p < .05$.

Results

Neuronal Necrosis

Diazepam and MK-801 reduced the severity of neuronal necrosis after exposure to soman relative to that of the animals given soman but with the test drug vehicle. The high and low doses of each test drug were equally effective in reducing necrosis. Therefore, the two dose groups were combined for each test drug in analyses of necrosis. Necrosis was virtually absent from all subjects in the soman-control groups.

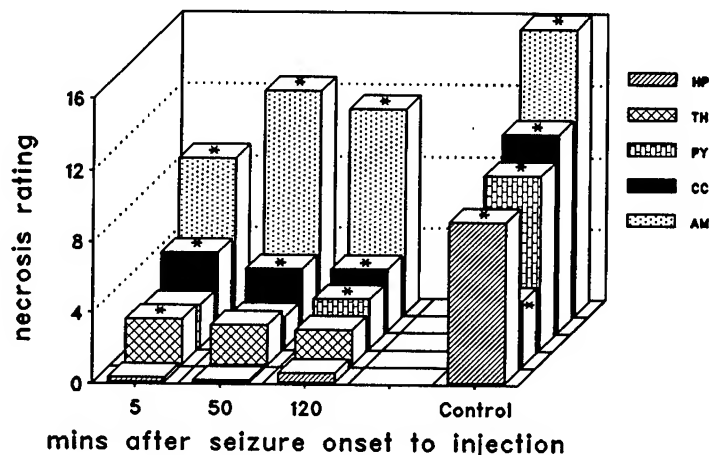


Figure 1. Diazepam moderately limited the development of necrosis equally at all delay times between EGSA appearance and drug administration. Average necrosis ratings by brain structure for groups treated with diazepam at 5, 50, or 120 min after the appearance of electrographic seizure activity. All subjects included here were challenged with $2 \times LD_{50}$ soman and given therapy drugs (see *Methods - Experimental Design* section). The low and high dose groups were combined for each injection delay time. Control = soman challenge plus combined diazepam/MK-801 vehicle groups (see text). Asterisk indicates significant difference ($p < .05$) from animals treated with diazepam but not challenged with soman. HP, hippocampus; TH, thalamus; PY, pyriform cortex; CC, cerebral cortex; AM, amygdala.

Diazepam limited the development of necrosis equally among the different injection delay groups poisoned with soman (see Fig. 1). There were no differences in necrosis levels within any brain area between

SPARENBORG, BRENNECKE, & BRAITMAN

groups receiving diazepam.

The hippocampus was protected by diazepam from soman-induced necrosis in all injection delay groups. The pyriform cortex was marginally protected, but the amygdala and cerebral cortex were not significantly protected. It is difficult to determine whether diazepam protected the thalamus, because this structure was only mildly damaged by soman in the test drug-control groups. The rating of the thalamus in the 5-min, but not the 50- or 120-min, injection delay group was significantly different from the diazepam/soman-control group. In the pyriform cortex, the rating for the 120-min, but not the 5- or 50-min, injection delay group was different from the diazepam/soman-control group.

In contrast, treatment with MK-801 at either 5 or 50 minutes after the appearance of EGSA gave virtually complete protection against neuronal necrosis (see Fig. 2). The level of necrosis was significantly greater in the subjects that received MK-801 at the 120-min injection delay than with either the 5- or 50-min delay. The level of necrosis in the thalamus, cerebral cortex, and amygdala of the subjects given MK-801 at the 120-min delay was significantly greater than in those subjects which did not receive soman.

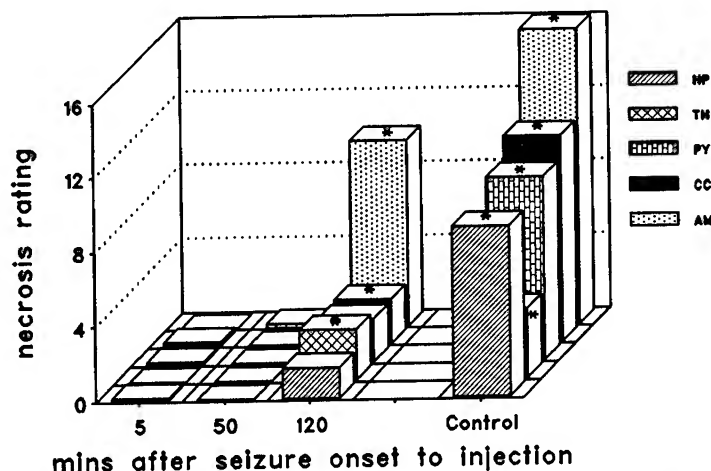


Figure 2. MK-801 completely prevented the development of necrosis in all brain structures when administered at less than 120 min after the appearance of electrographic seizure activity. The control group displayed here is the same as in Figure 1. Other information as in Figure 1.

In MK-801-control subjects, an abundance of necrotic neurons were found in each of the five brain areas examined. The amygdala was severely necrotic, but only mild damage occurred in the thalamus. The cerebral cortex, pyriform cortex and hippocampus all were moderately to severely damaged. Since there was only one survivor in the diazepam-control group, and the ratings for that subject were very similar to the group average for the MK-801-control group, that subject was added to the MK-801 group. The necrosis ratings for this combined group are found in Figs. 1 and 2.

SPARENBERG, BRENECKE, & BRAITMAN

Electrocorticogram data

Intense EGSA, indicative of status epilepticus, developed within 20 minutes of the soman injection in each test drug-control subject. EGSA intensity gradually declined during the six hour observation period, but a normal ECoG pattern did not return. Diazepam reduced EGSA soon after injection, but did not completely extinguish spike activity, resulting in very large z-scores in PSE data (see Fig. 3). In contrast, MK-801 either completely abolished or significantly reduced EGSA within 30 minutes of administration at each injection delay time as measured by PSE data (see Fig. 4).

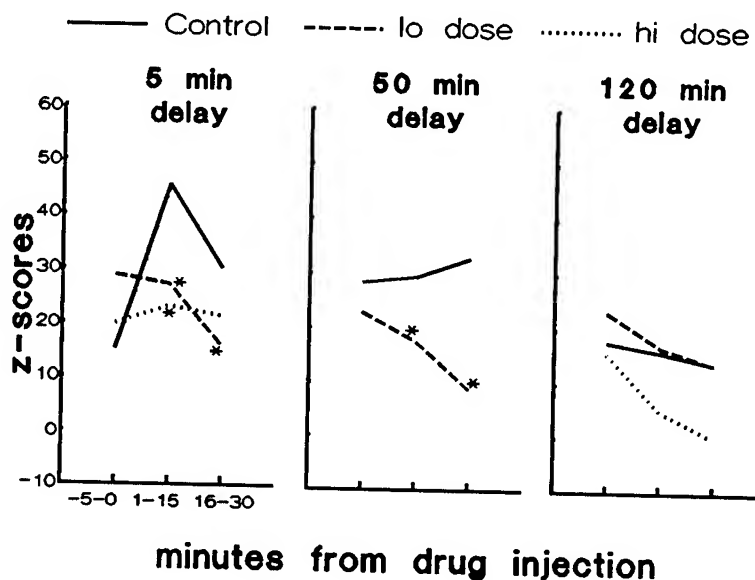


Figure 3. Diazepam reduced the intensity of electrographic seizure activity following its administration. Standardized power spectral estimates represent neurally generated electrical activity sampled for 5 min before the injection of diazepam (1 or 3 mg/kg) and during 2 15-min periods immediately following the injection. A z-score of zero represents normal brain activity recorded during the baseline period before any drug treatment. Asterisk indicates significant differences ($p < .05$) from the diazepam-control group at corresponding time periods.

In summary, both MK-801 and diazepam reduced the level of PSE after administration. In contrast to the effects of time of drug administration on necrosis, there was no indication that either drug had a time-dependent effect of reducing PSE. Generally, PSE from subjects treated with MK-801 were closer to zero, or normal, values than were diazepam-treated subjects.

SPARENBERG, BRENNER, & BRAITMAN

Convulsions

Each guinea pig developed severe motor convulsions which consisted of tonic limb extension mixed with rapid tonic-clonic limb movements. MK-801 effectively abolished motor convulsions at all doses and times of administration. Animals stopped convulsing and became prostrate and atonic following treatment with MK-801 independent of the delay in treatment. Convulsions were never entirely arrested by diazepam. Slight jerking movements of the head and/or whole body were evident for 48 hours after soman challenge in many survivors.

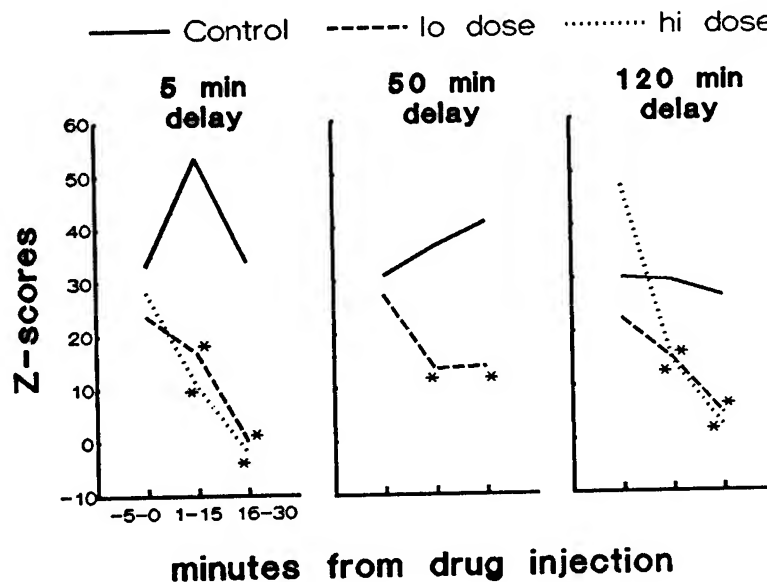


Figure 4. MK-801 (0.3 or 1 mg/kg) rapidly arrested or diminished EGSA when injected at any delay time. Other information as in Figure 3.

Recovery

All vehicle-treated subjects were unable to regain a normal stance after soman challenge. Their hindlimbs were spread abnormally and they did not use their hind feet to support themselves. Movement within their own cage was extremely limited. Subjects treated with MK-801 regained a normal stance and the ability to ambulate in their cages within 24 hours after the soman challenge. They were somewhat hyperactive at that time and were quick to avoid the caretaker's grasp. Their level of activity was near normal at 48 hours after soman. Diazepam-treated subjects,

SPARENBORG, BRENNECKE, & BRAITMAN

however, closely resembled the vehicle-treated subjects. Their stance was abnormal with spread hindlimbs and limited movement.

Lethality

Half of the subjects treated with soman plus the MK-801 vehicle did not survive for 48 hours after soman challenge, while only one of ten subjects in the soman plus diazepam vehicle group survived for 48 hours (see Table I). Survival in this latter group may have been adversely affected by the diazepam vehicle. However, the differences in survival rate between these groups were found not to be significant ($p > .05$).

MK-801 decreased lethality when given at 5 min after spike appearance, but not at longer delays. Lethality in all groups treated with diazepam was decreased relative to the diazepam SPV group, due to the high mortality rate in the diazepam SPV group.

Table I. Percent surviving at 48 hours after soman challenge.

Asterisk indicates significantly different from the test drug control group at $p < .05$ with the Fisher Exact test. Comparisons between test drug groups at corresponding doses and delay times were not significant.

Drug	MK-801			Diazepam		
	5	50	120	5	50	120
vehicle	43			10		
low dose	100*	70	40	80*	90*	60*
high dose	100*	nt	56	100*	nt	70*

nt indicates that no animals were tested under those conditions

Discussion

Treatment with MK-801 at 5 or 50 minutes after the appearance of electrographic seizure activity prevented the development of detectable neuronal necrosis within the five brain areas examined. In contrast, moderate to severe levels of necrosis were found in each diazepam-treated group.

SPARENBERG, BRENNECKE, & BRAITMAN

We have reported that MK-801 completely prevents brain damage when given to guinea pigs prior to soman challenge at doses (0.5 or 1 mg/kg) comparable to those used in this study (Braitman, Jaax & Sparenborg, 1988). Here, we have shown that MK-801 completely prevents brain damage when its administration has been delayed until nearly an hour after the onset of electrical seizures. In other work (Sparenborg, unpublished data) it was found that diazepam completely prevented neuronal necrosis when given to guinea pigs at a dose of 2 mg/kg prior to soman. However, in this experiment, when the administration of diazepam (3 mg/kg) was delayed until only 5 minutes after the onset of electrical seizures, a substantial level of necrosis developed. It appears that the efficacy of diazepam is decreased when given after a critical period early in the development of status epilepticus, but the efficacy of MK-801 was not decreased until the 120-minute injection delay time. The two drugs tested here apparently have different injection-delay-time limits for their ability to prevent or arrest necrosis. Such differences may have an important bearing on the medical protocol for controlling convulsions on the battlefield. Excitatory amino acid antagonists may be able to give higher levels of protection at longer delays after poisoning.

EAA antagonists such as MK-801 may have inherently superior efficacy against nerve agent-induced convulsions and brain damage than diazepam because EAA antagonists may directly antagonize the mechanisms which produce the pathology. Olney (1986) observed that the pattern of neuronal damage caused by convulsants such as pilocarpine and carbachol, which stimulate the acetylcholine (ACh) neurotransmitter system, as do nerve agents, is quite similar to the pattern of damage induced by other convulsants which directly stimulate the EAA neurotransmitter system. It may be that overstimulation of the ACh system by compounds such as carbachol and soman leads to overstimulation of the EAA system. The EAA neurotransmitter system is thought to be the causative system for such damage, rather than the ACh system, because the distribution of EAA, but not ACh, receptors is highly correlated with damaged brain structures.

Diazepam acts by facilitating general inhibition produced by the gamma-aminobutyric acid (GABA) neurotransmitter system, which may indirectly lead to the inhibition of EAA neurotransmitter release (Baba et al., 1983) and thereby indirectly protect against soman poisoning.

A desired capability of an anticonvulsant is that it provide for the rapid return to duty of victims of nerve agent poisoning. Although recovery of motor functions was not a focus of this experiment, our observations of guinea pig stance and mobility strongly suggest that a drug which antagonizes the EAA system, as MK-801 does, would provide faster return to duty than a benzodiazepine drug, such as diazepam.

The utility of an anticonvulsant will depend on the time of administration relative to nerve agent poisoning. Convulsions alone are capable of severely disrupting every normal function of the brain, including skeletal muscle control, respiration and cognition. Therefore,

SPARENBORG, BRENNECKE, & BRAITMAN

it is reasonable to expect that a drug capable of protecting the brain against convulsions would have side effects which would interfere with soldier performance if given in advance of a chemical attack. This problem necessitates post-attack administration of an anticonvulsant to those soldiers exhibiting early signs of nerve agent poisoning. The anticonvulsant must, therefore, be able to stop an ongoing convulsion. Since military units are required to quickly move from location to location on the modern battlefield, some stricken individuals may not receive aid until after the attack, requiring that an anticonvulsant be capable of arresting convulsions after the brain has experienced many minutes or hours of traumatic seizure activity.

Although diazepam did much to alleviate the toxicity of soman when administered after seizures began, it did not prevent necrosis completely. MK-801 did prevent all necrosis when injected at the 5- or 50-min delay times. If an anticonvulsant is needed on the battlefield but cannot be administered until after convulsions have begun, these results suggest that a drug which antagonizes the EAA system may offer a larger window of time in which complete protection against brain damage may be obtained. Given the efficacy of MK-801 in this and previous reports, drugs which interfere directly with the EAA system deserve serious consideration for future development as convulsant antidotes for nerve agents.

SPARENBORG, BRENNECKE, & BRAITMAN

REFERENCES

- Baba A, Okumura S, Mizuo H, Iwata H. Inhibition by diazepam and gamma-aminobutyric acid of depolarization release of cysteine sulfinic acid and glutamate in rat hippocampal slices. *J Neurochem* 1983; 40:280-284
- Braitman DJ, Jaax NK, Sparenborg SP. MK-801 protects against seizures and brain damage induced by the cholinesterase inhibitor soman. *Soc Neurosci Abstr* 1988; 14:240
- Braitman DJ, Sparenborg SP. MK-801 protects against seizures induced by the cholinesterase inhibitor soman. *Brain Res Bull* 1989; 23:145-148
- Clineschmidt BV, Williams M, Witoslawski JJ, Bunting PR, Risley EA, Totaro JA. Restoration of shock suppressed behavior by treatment with MK-801. *Drug Dev Res* 1982; 2:147-163
- DeVore JL. *Probability and statistics for engineering and the sciences*. Monterey, California. Brooks/Cole, 1982, pp. 596-598
- Foster AC, Gill R, Kemp JA, Woodruff GN. Systemic administration of MK-801 prevents N-methyl-D-aspartate-induced neuronal degeneration in rat brain. *Neurosci Lett* 1987; 76:301-311.
- Gilbert ME. The NMDA-receptor antagonist, MK-801, suppresses limbic kindling and kindled seizures. *Brain Res* 1988; 463:90-99
- Kochhar A, Zivin J, Lyden P, Mazzarella V. Glutamate antagonist therapy reduces ischemic damage. *Neurol* 1987; 37(Sup. 1):184
- Lemercier G, Carpenter P, Sentenac-Roumanau H, Morelis P. Histological and histochemical changes in the central nervous system of the rat poisoned by an irreversible anticholine-esterase organophosphorus compound. *Acta Neuropathol* 1983; 61:123-129.
- Lipp JA. Effect of benzodiazepine derivatives on soman induced seizure activity and convulsions in the monkey. *Arch Int Pharmacodyn* 1973; 202:244-251.
- McLeod CG, Singer AW, Harrington DG. Acute neuropathology in soman poisoned rats. *Fundam Appl Toxicol* 1984; 5:53-58.
- McNamara JO, Russell RD, Rigsbee L, Bonhaus DW. Anticonvulsant and antiepileptogenic actions of MK-801 in the kindling and electroshock models. *Neuropharm* 1988; 27:563-568
- Olney JW, Collins RC, Sloviter RS. Excitotoxic mechanisms of epileptic brain damage. *Adv Neurol* 1986; 44:857-878.
- Westerberg E, Kehr J, Ungerstedt U, Wieloch T. The NMDA-Antagonist MK-801 reduces extracellular amino acid levels during hypoglycemia and prevents striatal damage. *Neurosci Res Comm* 1988; 3:151-158
- Wong EHF, Kemp JA, Priestly T, Knight AR, Woodruff GN, Iverson LL. The anticonvulsant MK-801 is potent N-methyl-D-aspartate antagonist. *Proc Natl Acad Sci USA* 1986; 83:7104-7108

STEREOSPECIFICITY OF MICROBIAL ENZYMES FOR
G-AGENT DETOXIFICATION

Diane M. Steeves, Ms., John E. Walker, Mr.,
James J. Connolly, SP4, and Benedict J. Gallo, Dr.
U.S. Army Natick Research, Development and Engineering Center
Natick, Massachusetts 01760-5020

Francis C.G. Hoskin, Dr.
Illinois Institute of Technology
Chicago, Illinois 60616

INTRODUCTION

Enzymes from microbacteria capable of hydrolyzing a wide variety of organophosphates have been found both in procaryotic and eukaryotic organisms (1-7). In the past, these enzymes were called DFPases, somanases, paraoxonases or parathion hydrolases, based on their substrate specificity. The new name for these enzymes is organophosphorous acid (OPA) anhydases, which was selected during the First DFPase Workshop (Marine Biological Laboratory, Woods Hole, Massachusetts, June 1987) and relates to their basic activity. Fundamental questions pertaining to the molecular aspects of catalysis by these enzymes, such as the nature and sequence of chemical and perhaps other steps involved in organophosphate agent (OP) hydrolysis, character of the active site, and interactions between enzyme and substrate have been partially addressed by Dr. James Wild of Texas A and M (8-12), whose group has cloned the gene (*Opd*) which codes for an OPA Anhydrase that hydrolyzed diisopropyl fluorophosphate (DFP), isopropyl methylphosphonofluoridate (sarin) and 1,2,2-trimethylpropylmethylphosphonofluoridate (soman). A mechanism for a general base catalysis having a single in-line displacement at the phosphorous center with an activated water molecule was proposed. The enzyme has a molecular weight of 39,000 daltons, an isoelectric point of 8.3 and a requirement of zinc for reactivity. It is not known if the function of zinc is structural or catalytic. Mipafox was shown to be a competitive inhibitor.

An OPA anhydrase from the thermophile Bacillus stearothermophilus (JD-100) has been found to possess a different substrate specificity towards OP hydrolysis (5). The enzyme does not hydrolyze DFP but readily breaks down soman and sarin. The enzyme was manganous (Mn) ion requiring, as are most of the microbial OPA anhydases, but was not inhibited by mipafox. The enzyme can be covalently attached to cotton fabric with retention of its catalytic activity to hydrolyze soman (13). These properties have placed this enzyme in a separate category (Table 1) in the classification of OPA anhydases (14).

In this paper we will report on (i) the specificity of the OPA anhydrase from Bacillus stearothermophilus in hydrolyzing a series of OP agent surrogates [p-nitro-phenyl alkyl(phenyl)phosphinates]; (ii) the effect of cations alone and in combination on the enzyme's reactivity to the OP surrogate and (iii) the ability of the enzyme to detoxify soman. A comparison will be made with other microbial OPA anhydases (E.coli, Tetrahymena thermophila, Flavobacterium B4079) and also with the OPA anhydrase from squid hepatopancreas.

Approved for public release
Distribution unlimited.

TABLE 1

CHARACTERISTICS OF OPA ANHYDRASES

SQUID TYPE
CEPHALOPOD NERVE, HEPATOPANCREAS, AND SALIVA ONLY
MOLECULAR WEIGHT, 30,000
HYDROLYZES DFP 4-5 TIMES FASTER THAN SOMAN AT 3 mM
Mn ²⁺ INDIFFERENT OR SLIGHTLY INHIBITED
STRONGLY INHIBITED BY 10 ⁻⁴ M EDTA, EGTA; INDIFFERENT TO 1, 10-PHENANTHROLINE AND 8-HYDROXYQUINOLINE-5-SULFONATE
STABLE IN 80-85% SATURATED AMMONIUM SULFATE
MIPAFIX INDIFFERENT
HYDROLYZES ALL ISOMERS OF SOMAN; SOME STEREOSPECIFICITY
OT TYPE
ONE SUBSPECIES OF <i>BACILLUS STEAROTHERMOPHILUS</i>
MOLECULAR WEIGHT, 82-84,000
HYDROLYZES SOMAN; DOES NOT HYDROLYZE DFP AT ALL
Mn ²⁺ STIMULATED 80-FOLD AT 10 ⁻⁴ M
STRONGLY INHIBITED BY EDTA, EGTA, 1, 10-PHENANTHROLINE, AND 8-HYDROXYQUINOLINE-5-SULFONATE
STABLE IN AMMONIUM SULFATE
MIPAFIX INDIFFERENT
HYDROLYZES ALL ISOMERS OF SOMAN
MAZUR TYPE
UBIQUITOUS; EVEN IN OTHER CEPHALOPOD TISSUES
MOLECULAR WEIGHTS VARY, 45-90,000; PROBABLY SPECIES SPECIFIC
HYDROLYZES SOMAN 5-50 TIMES FASTER THAN DFP AT 3 mM
Mn ²⁺ STIMULATED 2- TO 20-FOLD
AMMONIUM SULFATE LABILE
MIPAFIX INHIBITED

MATERIALS AND METHODS

Culture and Growth Conditions of the Test Microorganisms.

1. Bacillus stearothermophilus strain I5B: The original strain (JD-100) was supplied by Dr. Joseph DeFrank, Chemical Research, Development and Engineering Center, Edgewood, Maryland. It was grown as a submerged culture in multiple 2800 mL Fernbach flasks containing 500 mL of LB broth (15) with streptomycin. They were started with 50 mL of 24-hour-old seed cultures, shaken at 75 to 80 strokes per minute at 55°C and usually harvested after 24 hours.
2. Flavobacterium SP. strain B4079: The strain was grown as a submerged culture using Cook's medium (16). The cultures were shaken in Fernbach flasks at 29°C.
3. Tetrahymena thermophila, strain "brynat": The culture was grown in a 14 liter stirred tank reactor (fermentor) with a working volume of 10 liters using Modified Cook's Medium (16) with chelated iron and streptomycin, at 29°C.
4. E.coli, strain ATCC25922: The culture was grown in a stirred fermentor using LB Basic Broth (15) at 37°C.

Extraction of OPA Anhydrase from Cultured Cells.

The harvested cells were concentrated by centrifugation and washed two times with cold sonicate buffer (400 mM potassium chloride (KCL), 50 mM sodium chloride (NaCL) and 5 mM in 1,3-bis(tris(hydroxymethyl)-methylamino) propane (Bis-Tris-Propane), pH7.2. A 20% (w/v suspension of packed cells in the sonicate buffer was sonicated with a 19 mm probe, 80% output, 50% duty, in 10 one-minute periods alternating with one-minute rest periods. The temperature was held at or below 10°C with an ice bath. The sonicate was centrifuged at 90,000 X g and the clarified supernatant used as the source material for further enzyme purification.

Assays of OPA Anhydrase from Cultured Cells.

1. Synthesis of substrates: 4-nitrophenyl methyl(phenyl)phosphinate (PNMPP) was synthesized by the Method of Horton (17). 4-nitrophenyl ethyl(phenyl)phosphinate (PNPEPP) was obtained commercially (Ash Stevens, Inc.) and 4-nitrophenyl isopropyl(phenyl)phosphinate (PNPiPP) was a gift from E.Durst of CRDEC. Soman was synthesized essentially according to the Method of Larsson (18). Diisopropylfluorophosphate (DFP) was obtained commercially (Sigma).

2. Assay procedures: In the typical assay, 50 microliters (uL) of a solution containing the OPA anhydrase was added to 2.5 mL of sonicate buffer. The hydrolysis reaction was initiated by the addition of 50 uL of a 3×10^{-3} M solution of the 4-nitrophenyl phosphinate substrate (PNMPP, PNPEPP or PNPiPP) in acetonitrile. The rate of hydrolysis of the substrate was monitored by the release of p-nitrophenyl anion as measured by absorbance at 400nm. Quantitation was made by reference to a calibration curve of known concentrations of p-nitrophenol released per minute. The hydrolysis of the fluorophosphonate substrated Soman and DFP were monitored with a fluoride-sensitive electrode after the method of Hoskin (19). Determination of the loss of acetylcholinesterase (AChE) inhibitory activity was performed as described by Hoskin (20).

Pseudo first order rate constants (k') were obtained by fitting the hydrolysis rate data to the equation: $A_t = A + (A_0 - A)e^{-k't}$ where

A_0 , A_t and A are the p-nitrophenolate anion concentrations at times zero, t , and infinity. Best fit values of k' and A were obtained using a computerized least squares method (KORE).

Purification of OPA Anhydrase from Cultured Cells.

The supernatant from the high speed centrifugation of the cultured cells was concentrated by ultrafiltration over a PM-10 membrane (Amicon Corp). The concentrated supernatant was applied to a G150 sephadex (Pharmacia) column 2.5 x 90 cm and eluted with the sonicate buffer. The protein content of the effluent was monitored by absorbance at 280nm and the OPA anhydrase was measured by the PNPMP assay method. The chromatographic fractions exhibiting OPA anhydrase activity were pooled and dialyzed against Tris buffer, pH7.5. The fraction was applied to a DEAE A-50 (Pharmacia) column, 1.5 x 30 cm. The OPA anhydrase was eluted after the breakthrough peak by means of a linear gradient with a limit buffer of 20 mM Tris, 0.5 MNaCl, pH7.5. In the final stage of purification, the DEAE gradient fraction containing OPA anhydrase activity after dialysis against 20 mM Tris, pH7.5 was applied to a 1 x 50 cm column of Phenyl-CL Sepharose 4B (Pharmacia) previously equilibrated in 1.7 M ammonium sulfate, $(\text{NH}_4)_2 \text{SO}_4$ in 20 mM Tris pH7.5. The column was developed with a linear gradient of 400 mL using a starting buffer of 1.7M $(\text{NH}_4)_2 \text{SO}_4$ in 20 mM Tris pH7.5 and a limit buffer of 20 mM Tris, pH7.5. The purified enzymes were assayed for activity against the p-nitrophenylalkylphenylphosphinates as well as against soman and DFP. Protein concentration was determined by the Coomassie dye binding assay of Bradford (21). The molecular weights of the purified OPA anhydases were determined by SDS polyacrylamide gel electrophoresis using the discontinuous buffer method of Laemmli (22). The OPA anhydases were also run under nondenaturing conditions using linear gradient precast polyacrylamide gels (Jule Inc.).

Addition of Metal ions and Metal ion Chelators to OPA anhydrase.

Metal ions and chelators were made up in 25mM HEPES buffer pH7.2 at 100mM concentrations. Suitable additions were made to the enzyme solutions to give a range of concentrations from 1mM to 0.1mM. The reaction proceeded at room temperature for 30 minutes and changes in OPA anhydrase activity were monitored by the PNPMP assay method.

RESULTS

The separation of the OPA anhydrase from Bacillus stearothermophilus I5B was performed by gel filtration with G150 (Figure 1) followed by hydrophobic chromatography using phenyl-sepharose CL-4B (Figure 2).

FIGURE 1
CHROMATOGRAPHIC SEPARATION OF I5B ON G-150
15B G150-CHROMATOGRAPHY
PROTEIN & ACTIVITY CURVES

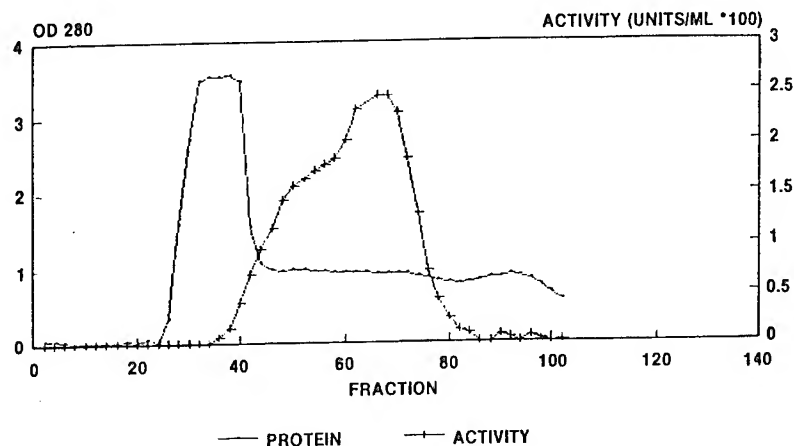
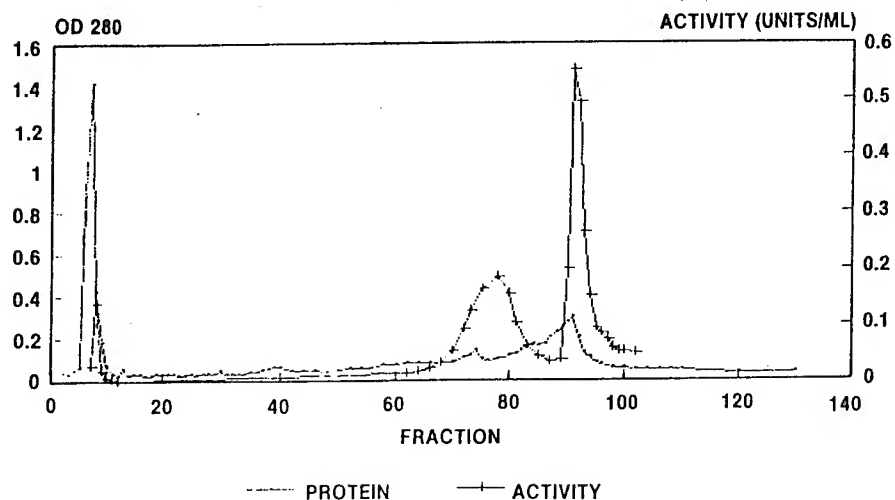


FIGURE 2
CHROMATOGRAPHIC SEPARATION OF I5B ON PHENYL-SEPHAROSE CL-4B
PHENYL-SEPHAROSE CHROMATOGRAPHY OF IB5
PROTEIN & ACTIVITY CURVES



This separation methodology was followed for the enzyme isolation from the other microbial organisms. All the OPA anhydases appeared to have a high degree of hydrophobicity since their separation from the phenyl-sepharose column was at the end of the elution gradient in every case. The molecular weights of the OPA anhydases are shown in Table 2.

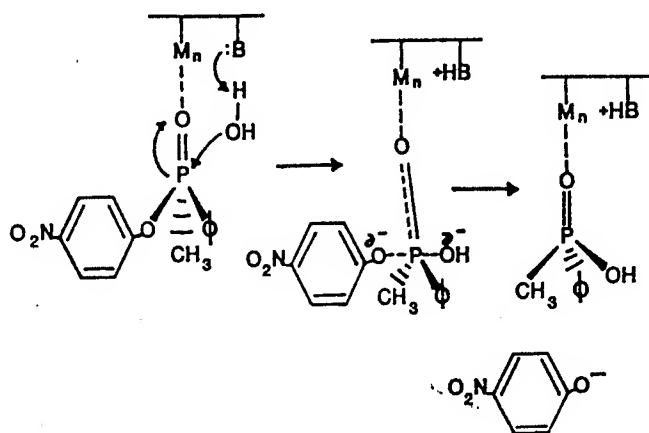
TABLE 2
MOLECULAR WEIGHTS OF THE OPA ANHYDRASES BY SDS ELECTROPHORESIS

MOLECULAR WEIGHTS OF MICROBIAL OPA ANHYDRASES

SOURCE OF OPA ANHYDRASE	MOLECULAR WEIGHT (DALTONS)		
<i>Bacillus stearothermophilus</i>	62,000	40,000	36,000
<i>Flavobacterium</i> SP 4079	62,000	58,000	
<i>Tetrahymena Thermophila</i>	45,000		
<i>E. coli</i> ATCC 25922			
F-1	53,000		
F-2	38,000		

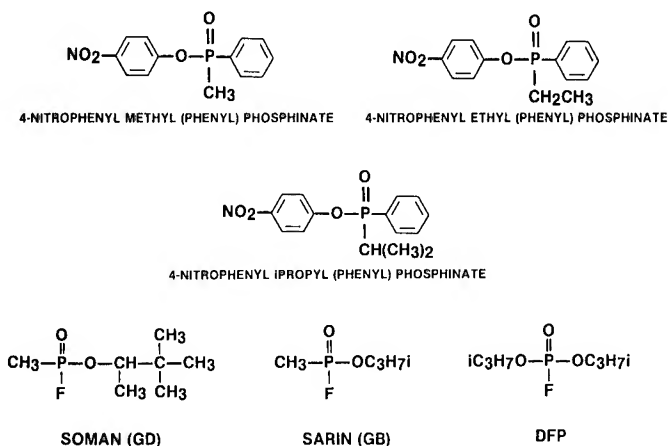
A proposed model for the enzymatic hydrolysis of p-nitrophenyl methyl (phenyl)phosphinate by microbial OPA anhydases is shown in Figure 3.

FIGURE 3
MECHANISM FOR HYDROLYSIS OF PNPMP BY MICROBIAL OPA ANHYDRASE



The model, which is based on a similar model for phosphotriesterase (II), postulates a single base with a Pk_a of 6.1 in the enzyme's active site that abstracts a proton from water. The activated water molecule attacks directly at the phosphorous center with the overall stereochemical outcome of inversion of configuration. Table 3 shows the organophosphorous AChE inhibitors, which were used as substrates in the study.

TABLE 3
ORGANOPHOSPHOROUS AChE INHIBITORS



The initial and pseudo-first order rate constants for the enzymatic hydrolysis of p-nitrophenyl alkyl(phenyl)phosphinates are given in Table 4.

TABLE 4
INITIAL AND PSEUDO-FIRST ORDER RATE CONSTANTS FOR THE HYDROLYSIS
OF P-NITROPHENYL AKLYL(PHENYL) PHOSPHINATES

ENZYME SOURCE		PNMPMP	PNPEPP	PNPIPP
BACILLUS STEAROTHERMOPHILUS				
JD-100				
	k'	3.77E-01	4.50E-02	N.R
	K'	2.40E-03	2.80E-04	N.R
JD-200				
	k'	3.32E-01	3.00E-02	N.R
	K'	2.10E-03	2.20E-04	N.R
JD-300				
	k'	1.81E-01	N.R	N.R
	K'	1.30E-03	1.50E-04	N.R
15B				
	F-1 k'	5.40E-02	1.20E-02	N.R
	K'	1.70E-03	2.10E-04	N.R
	F-2 k'	1.51E-01	1.80E-02	N.R
	K'	4.00E-03	8.10E-04	N.R
B4079 FLAVOBACTERIUM				
	F-1 k'	1.57E-01	9.00E-02	6.00E-03
	K'	4.80E-03	3.80E-03	1.90E-04
	F-2 k'	5.40E-02	2.40E-02	6.00E-03
	K	1.90E-03	1.26E-03	1.50E-04
ATCC25922 E. COLI				
	F-1 k'	2.70E-02	1.50E-02	3.00E-03
	K'	1.60E-03	9.50E-04	1.20E-04
	F-2 k'	1.80E-02	3.00E-03	N.R
	K'	1.00E-03	4.30E-04	5.00E-05
SQUID OPA ANHYDRASE				
	k'	4.30E-02	5.00E-03	3.00E-03
	K'	3.00E-03	9.10E-04	3.80E-05

The effect of metal cations on the reactivity of microbial OPA anhydases is shown in Table 5.

TABLE 5
RECOVERY OF OPA ANHYDRASE ACTIVITY
AFTER CHELATION INHIBITION

ENZYME SOURCE	Mn2 +	Mg2 +	Ca2 +	Mg2 +/ Mn2 +	Ca2 +/ Mn2 +
BACILLUS STEAROTHERMOPHILUS					
JD-100	+	-	-	+/-	+/-
JD-200	+	-	-	+/-	+/-
JD-300	+	-	-	+/-	+/-
15B					
F-1	+/-	-	-	+/-	+/-
F-2	+/-	-	-	+/-	+/-
B4079 FLAVOBACTERIUM					
F-1	-	-	-	-	-
F-2	+	-	-	+	+
E. COLI					
F-1	+	-	-	+	+
F-2	+	-	-	+	+
SQUID					
	+	-	-	+	+

The ability of the microbial OPA anhydases to hydrolyze and detoxify soman is shown in Table 6.

TABLE 6
RELATIVE RATES OF HYDROLYSIS OF SOMAN BY MICROBIAL OPA ANHYDRASES

CHARACTERIZATION OF ENZYME SAMPLES WITH RESPECT TO THEIR ACTIVITY (HYDROLYSIS AND DETOXIFICATION) AGAINST SOMAN			
SAMPLES		PERCENT ACTIVITY	
		HYDROLYSIS	DETOXIFICATION
<i>Bacillus stearothermophilus</i>	JD-100	53	49
	JD-200	40	13
	I5B	53	45
Tetrahymena		50	39
Flavobacterium		47	7
E. coli FI		50	50
E. coli FII		55	20

DISCUSSION

All of the OPA anhydases from the microbacteria eluted from the hydrophobic chromatographic separation step at the end of the elution gradient similarly to the I5B separation shown in Figure 2. This fact indicates that these proteins tended to possess a high degree of hydrophobicity on their exterior structure. Since sonication was needed to release them from the microorganisms, they are probably compartmentalized in the bacterial cell wall. The OPA anhydases from *Bacillus stearothermophilus* showed discrimination between the methyl and ethyl homologues of the phenyl phosphinates but the OPA anhydases from the other strains hydrolyzed the methyl or ethyl derivatives at essentially the same rate as seen in Table 4. All of the enzymes were manganous ion requiring, as is shown in Table 5. Magnesium or calcium ions could not reactivate the activity of the enzyme after EDTA inhibition. There was no evidence that calcium or magnesium bound to the manganous ion binding site in the OPA anhydases. It was found that the level of EDTA concentration to inhibit the activity of *E.coli* OPA anhydrase was an order of magnitude higher than was necessary to inhibit the other enzymes, which may indicate cofactor status for Mn^{+2} in the active site of the *E.coli* enzyme. Table 6 shows that of the OPA anhydases tested for their ability to both hydrolyze and detoxify soman, JD-200, B4079 and f_{II} of *E.coli* preferentially hydrolyzed the nontoxic soman isomers. This differential activity was not predicted from initial

and pseudo first order rate constants shown in Table 4. The 4-nitrophenyl alkyl(phenyl)phosphinates were racemic mixtures and this probably prevented a discrimination in their reactivity to the microbial OPA anhydrases.

CONCLUSIONS

1. OPA anhydrases from microbacteria possess a high order of hydrophobicity.
2. OPA anhydrase from Bacillus stearothermophilus was the most selective in the hydrolysis of the soman surrogates.
3. The OPA anhydrase from E. coli may require a specific binding site for manganous ion for reactivity.

REFERENCES

1. Mazur, A. An Enzyme in Amlinal Tissue Capable of Hydrolyzing the Phosphorous-fluorine Bond of Alkyl Fluorophosphates. J. Biol. Chem. 164:271 (1946).
2. Mounter, L.A., Baxter, R.F. and Chautin, A. Dialkylfluorophosphatases of Microorganisms. J. Biol. Chem. 215:699 (1955).
3. Zech, R. and Wigand, K.D. Organophosphate-detoxifying Enzymes in E. coli, Gel Filtration and Isoelectric Focusing of DFPase, Paroxonase and Unspecified Phosphohydrolases. Experimentia 31:157 (1975).
4. Landis, W.G., Savage R.E. Jr., and Hoskin, F.C.G. An Organofluorophosphate-hydrolyzing Activity in Tetrahymena Thermophilia. J. Protozool. 32:517 (1985).
5. Chettur, G., DeFrank, J.J., Gallo, B.J., Hoskin, F.C.G., Mainer, S., Robbins, F.M., Steinmann, K.E. and Walker, J.E. Soman Hydrolyzing and Detoxifying Properties of an Enzyme From a Thermophillic Bacterium. Fund. Appl. Tox. 11:127 (1988).

This paper reports research sponsored by US Army Natick Research, Development and Engineering Center and has been assigned TP-2837 in the series of papers approved for publication.

REFERENCES (CONT'D)

6. Cheng, F.C., Deas, R.A., DeFrank, J.J. and Elashvili, I. Purification Procedures for OPA Anhydrase-2 From Halophile JD 6.5. 1989 U.S. Army Chemical Research, Development and Engineering Center Scientific Conference on Chemical Defense Research, Aberdeen Proving Ground, Maryland.
7. Dumas, D.P., Wild, J.R. and Raushel, F.M. Diisopropylfluorophosphate Hydrolysis by an Organophosphate Anhydrase From Pseudomonas diminuta. Biotech. and Appl. Biochem. 11:235 (1989).
8. McDaniel, C.S., Harper, L.L. and Wild, J.R. Cloning and Sequencing of a Plasmid-Borne Gene (opd) Encoding a Phosphotriesterase. J. Bact. 170:2306 (1988).
9. Lewis, V.E., Donarski, W.J., Wild, J.R. and Raushel, F.M. Mechanism and Stereochemical Course at Phosphorous of the Reaction Catalyzed by a Bacterial Phosphotriesterase. Bioch. 27:1591 (1988).
10. Donarski, W.J., Dumas, D.P., Heitmeyer, D.P., Lewis, V.E. and Raushel, F.M. Structure-Activity Relationships in the Hydrolysis of Substrates by the Phosphotriesterase From Pseudomonas diminuta ibid. 28:4650 (1989).
11. Dumas, D.P., Caldwell, S.R., Wild, J.R., and Raushel, F.M. Purification and Properties of the Phosphotriesterase from Pseudomonas diminuta J. Biol. Chem. 264:19659 (1989).
12. Dumas, D.P., Durst, H.D., Landis, W.G., Raushel, F.M. and Wild, J.R. Inactivation of Organophosphorous Nerve Agents by the Phosphotriesterase from Pseudomonas diminuta Arch. Biochem. Biophys. 277:155 (1990).
13. Walker, J.E., Gallo, B.J., Starcher, B., Robbins, F.M., and Hoskin, F.C.G. Soman Hydrolyzing and Detoxifying Properties of an Enzyme from a Thermophilic Bacterium. 1988 Army Science Conference, Fort Monroe, Virginia.
14. Hoskin, F.C.G., Kirkish, M.A., and Steinmann, K.E. Two Enzymes for the Detoxification of Organophosphorous Compounds-Sources, Similarities and Significance. Fund. Appl. Tox. 4:S165 (1984).

REFERENCES (CONT'D)

15. Lennox, E.S. Transduction of Linked Genetic Characters of the Host by Bacteriophage P1. *Virology* 1:190 (1955).
16. Cook, A.M. and Brown, M.R.W. The Relationship Between Heat Activation and Colony Formation for the Spores of Bacillus stearothermophilus. *J. Pharm. Pharmacol.* 16:725 (1964).
17. Horton, G.L., Lieske, C.N., and Lowe, J.R. Phosphinate Inhibition Studies of Cholinesterases. *Pestic. Sci.* 9:135 (1978).
18. Larsson, L. Alkaline Hydrolysis of Isopropoxymethyl-phosphoryl fluoride (Sarin) and Some Analogues. *Acta. Chem. Scand.* 11:1131 (1957).
19. Hoskin, F.C.G. Inhibition of a Soman and Diisopropylphosphorofluoridate (DFP) Hydrolyzing Enzyme by Mipafox. *Biochem. Pharmacol.* 34:2069 (1985).
20. *ibid.* An Organophosphorous Detoxifying Enzyme Unique to Squid in Squid as Experimental Animals; Gilbert, D.L., Adelman, W.J. Jr., and Arnold, J.M., editors, Plenum, New York, Chap. 21, 1990.
21. Bradford, M. A Rapid Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding. *Anal. BioCh.* 72:248 (1976).
22. Laemmli, U.K. Cleavage of Structure 1 Proteins During the Assembly of the Head of Bacteriophage T4. *Nature* 227:680 (1970).

STEINHAUS, BASKIN, CLARK & KIRBY

Formation of Methemoglobin and Metmyoglobin Using
8-Aminoquinoline Derivatives or Sodium Nitrite and
Subsequent Reaction With Cyanide

Ralph K. Steinhaus, Ph.D.

*Steven I. Baskin, Pharm.D., Ph.D.

James H. Clark

Stephen D. Kirby

United States Army Medical Research Institute of Chemical Defense
Aberdeen Proving Ground, MD 21010-5425

INTRODUCTION

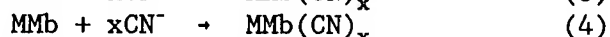
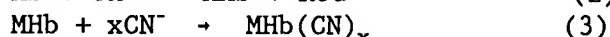
The first step in a common method for the treatment of cyanide poisoning is to induce methemoglobinemia by any of a variety of oxidants¹. The reaction of cyanide with methemoglobin (MHb), but not hemoglobin (Hb)², provides a pathway for the removal of cyanide, thus preventing it from reacting with cytochrome oxidase and thereby inhibiting aerobic respiration³. However, Hb is not the only heme protein that undergoes oxidation. Myoglobin (Mb), an intracellular protein, both transports and stores oxygen in red skeletal muscle fibers and is also oxidized by the same agents that oxidize Hb⁴.

A number of studies of Hb oxidation have been performed in vivo, yielding qualitative data but not rate constants. The exception to this is the reaction between Hb and NO_2^- which has been extensively studied⁵⁻¹¹. Recently, a series of 8-aminoquinoline derivatives have been shown to be effective MHb formers in dogs¹². Obtaining quantitative kinetic information on the reaction of these 8-aminoquinoline derivatives with Hb in vitro should provide insight as to the mechanisms of these agents and how to rationally employ them under in vivo conditions. Further, virtually no information exists on the reaction of cyanide with MHb in the presence of oxidants. It has been shown that excess nitrite ion (NO_2^-), an oxidant for heme proteins, does bind reversibly to MHb^{13,14} and metmyoglobin (MMb)¹⁵, although NO_2^- is far less tightly bound than cyanide ion and is readily displaced by cyanide ion¹⁶. Nevertheless, the possibility exists that certain oxidants could affect the binding of cyanide ion by MHb.

It was the purpose of this study to investigate the kinetics of the oxidation of Hb and Mb in vitro using the 8-aminoquinoline derivatives

STEINHAUS, BASKIN, CLARK, & KIRBY

listed below and NO_2^- as a control. The kinetics of the binding of cyanide ion to MHb and MMb in the presence of oxidant were also studied to determine if any interference resulted from the presence of the oxidant. The oxidants used in this study were 8-[(4-amino-1-methylbutyl)amino]-6-methoxy-quinoline di-phosphate (primaquine), 6-methoxy-8-(6-dimethylaminohexylamino)-4-methyl-quinoline dihydrochloride (WR6026), 8-[(4-amino-1-methylbutyl)amino]-2,6-dimethoxy-4-methyl-5-[(3-trifluoromethyl)phenoxy]quinoline succinate (WR238,605), and sodium nitrite. The reactions followed are shown in equations 1 - 4.



Ox and Red indicate the oxidized and reduced forms respectively, of oxidant, CN^- indicates cyanide ion, and $\text{MHb}(\text{CN})_x$ and $\text{MMb}(\text{CN})_x$ represent cyanomethemoglobin and cyanometmyoglobin, respectively. The other symbols, Mb and Hb were previously defined.

MATERIALS AND METHODS

Reagents: All compounds were reagent grade unless otherwise specified. Hb was obtained by water-lysing outdated human red blood cells followed by centrifugation at 15,000 x g. The Hb was stored at 4°C. Samples of oxyhemoglobin were freshly prepared as needed. Aliquots of Hb were saturated with oxygen¹⁶ at 5-10°C and treated with a 100-fold excess of sodium dithionate. The excess dithionate was removed by passing the sample through a Sephadex 25-G column whose bed volume measured 2.0 by 30 cm. The samples appeared cloudy at that point and were filtered through a 0.1 micrometer Millipore filter. The solvent was 0.05 M N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES), adjusted to pH 7.4. Dog Mb was obtained from Sigma (St. Louis) and stored at 4°C. MHb was prepared as needed and worked up following the same procedure as was used for Hb, except that filtration was not necessary. Primaquine, WR6026, and WR238,605 were obtained from the inventory of Walter Reed Army Institute of Research and were of reagent grade quality.

Hb and Mb concentrations were determined by conversion to cyanomethemoglobin and cyanometmyoglobin, respectively, using $\text{K}_3\text{Fe}(\text{CN})_6$ and KCN with subsequent measurement of the absorbance at 540 nm. The accepted standard value of the molar absorptivity of cyanomethemoglobin was used and is 4.4×10^4 ¹⁷, while that of cyanometmyoglobin was measured and found to be 9.83×10^3 . All other compound concentrations were determined by weight. HEPES (0.05 M), adjusted to pH 7.4, was used to prepare all solutions except those of WR238,605, which were prepared using DMSO.

Spectra and absorbance measurements: They were obtained on a Cary Model

UNCLASSIFIED

STEINHAUS, BASKIN, CLARK, & KIRBY

118 recording spectrophotometer. Oxidation reactions were followed on a Beckman DU-7 spectrophotometer with a kinetics attachment, while the cyanide complexation reactions required the use of a Hi-Tech Scientific Stopped-Flow spectrophotometer.

Methodology: Concentration effect curves were obtained by mixing together a fixed amount of Hb or Mb and varying amounts of oxidant. The absorbance of the Hb solutions was measured at 630 nm against a blank solution containing Hb and no oxidant. At this wavelength, only MHb absorbs. The literature value of the molar absorptivity of MHb, 1.52×10^4 l⁸, was used to calculate the MHb concentrations. The absorbance of the Mb solutions were measured at 580 nm against a buffer blank. At this wavelength, both Mb and MMb absorb. The molar absorptivities used to calculate the concentrations of Mb and MMb were measured at 580 nm and found to be 1.13×10^4 and 3.16×10^3 , respectively. All solutions were kept in a water bath at 37°C.

Reaction rates for the oxidation of Hb to MHb were followed at 630 nm. Solutions were pre-equilibrated at 37°C. The reactions were initiated by adding a measured volume of oxidant to a measured volume of Hb, already in the spectrophotometer cell; cell temperature was maintained at 37°C. The same procedure was followed for measurement of the oxidation rates of Mb, except that the wavelength was 580 nm. The rates of reaction of cyanide ion with MHb and MMb were rapid and required the use of a stopped-flow spectrophotometer. The reactions were followed at 540 nm. MHb and MMb were formed using a 10-fold excess of oxidant. The excess oxidant and its reduced form were not removed from the solution. Solutions were pre-equilibrated at 37°C, and the storage syringe and cell temperature were each maintained at 37°C.

RESULTS

Concentration Effect Curves: Concentration effect curves for the oxidation of Hb could only be obtained within 45 min. after mixing because, within one to two hours of mixing, precipitation occurred in the solutions. In the case of WR238,605, precipitation occurred almost immediately and no data were obtained. Standard deviations for all concentration effect curve points were $\pm 3\%$ or less. Figure 1 shows that there is a linear relationship between the MHb formed and the amount of oxidant added.

Concentration effect curves for the oxidation of Mb were obtained during the full time span of the reaction. No data were obtained for the reaction of WR238,605 due to precipitation. Figure 2 illustrates the rate of oxidation at two different times during the course of the reaction.

Oxidation Kinetics: Hb: The oxidation of Hb to MHb using primaquine, WR6026, or WR238,605 as the oxidant proceeded with the formation of a

STEINHAUS, BASKIN, CLARK, & KIRBY

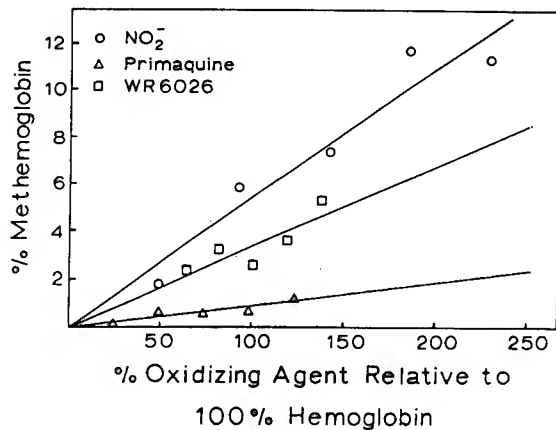


Figure 1. Concentration effect curve for formation of methemoglobin using NO_2^- , primaquine, and WR6026 at 37°C and $\text{pH} = 7.4$. Data taken 45 min. after reactions were initiated.

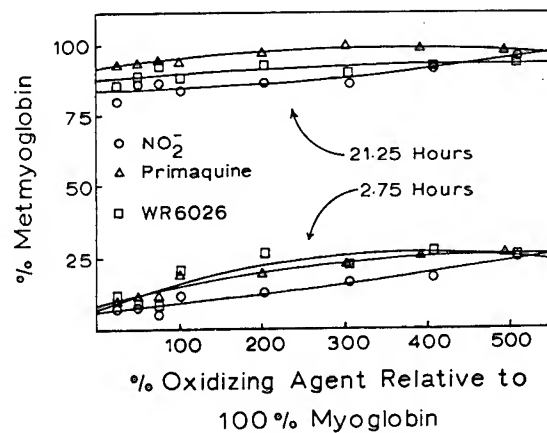


Figure 2. Concentration effect curve for formation of metmyoglobin using NO_2^- , primaquine, and WR6026 at 37°C and $\text{pH} = 7.4$. Data taken at both 2.75 and 21.25 hours after reactions were initiated.

precipitate, making it impossible to obtain kinetic data for the oxidation reactions. An Hb concentration of 7.25×10^{-6} M in the presence of a 10-fold molar excess of WR6026 was sufficient to cause precipitation which suggests solubility limits lower than concentrations employed. At this concentration, the absorbance change due to product formation and precipitation was very slight.

The oxidation of Hb by sodium nitrite proceeded without precipitation during the course of the reaction. The reactions had a lag phase that occurred prior to the oxidation reaction as shown in Figure 3. The autocatalytic reaction occurring after the lag phase was shown (spectral example, Figure 3) to be first-order in Hb and obeyed the first-order rate equation

$$\frac{-d[\text{heme protein}]}{dt} = k_o [\text{heme protein}] \quad (5)$$

where k_o represents the pseudo-first-order rate constant and is described in equation 6. The autocatalytic portion of the reaction has been shown to account for 90% of the total oxidation¹⁰. Therefore, only data points after the first half-life were used so that no portion of the lag phase could influence the data.

$$k_o = k \frac{\text{NO}_2^-}{\text{heme protein}} [\text{NO}_2^-] \quad (6)$$

STEINHAUS, BASKIN, CLARK, & KIRBY

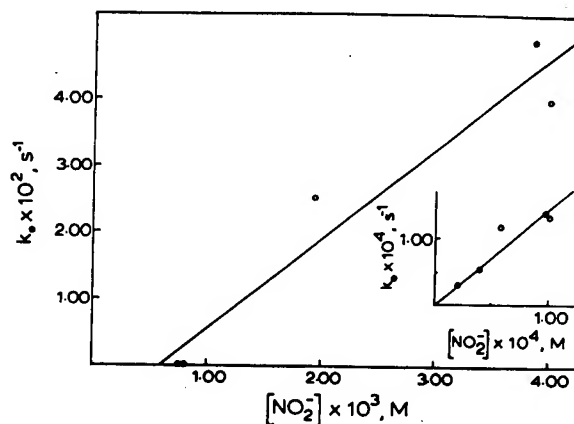
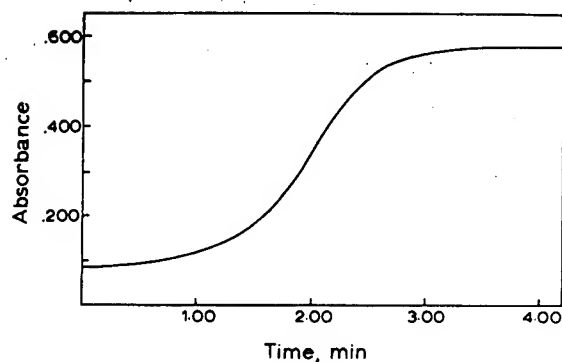
The reactions were carried out with a 10- to 50-fold molar excess of nitrite. The data yielded excellent first-order plots, thus demonstrating the validity of equation 5. A plot of the pseudo-first-order rate constant vs. nitrite ion concentration was linear over a concentration range (see Figure 4 including insert), thus establishing the order in nitrite as first and allowing resolution of the second-order rate constant, $k_{\text{Hb}}^{\text{NO}_2^-}$. The intercept of the plot was negative, suggesting a shift in rate constant at nitrite ion concentrations lower than those used in this study. The value of the second-order rate constant, listed in Table 1, was $74 \text{ M}^{-1}\text{s}^{-1}$ over the nitrite ion concentration range of $4.3 \times 10^{-4} \text{ M}$ to $2.2 \times 10^{-3} \text{ M}$.

Table 1

Resolved Rate Constants for The Oxidation of Hb and Mb
by Various Oxidizing Agents. Std. Dev. are Given.

Hb		
Oxidizing Agent	pH	$k_{\text{Hb}}^{\text{Ox}}, \text{M}^{-1} \text{s}^{-1}$
NO_2^-	7.53	74.1 ± 4.6

Mb			
Oxidizing Agent	pH	$k_{\text{Mb}}^{\text{Ox}}, \text{M}^{-1} \text{s}^{-1}$	$k_{\text{Mb}}, \text{s}^{-1}$
NO_2^- , low conc	7.45	$(1.79 \pm 0.03) \times 10^{-1}$	$(2.75 \pm 0.18) \times 10^{-5}$ $(1.37 \pm 0.02) \times 10^{-4}$
NO_2^- , high conc	7.61	13.4 ± 3.0	
WR6026	7.50	$(5.15 \pm 0.73) \times 10^{-3}$	
Primaquine	7.48	$(2.88 \pm 0.09) \times 10^{-2}$	



STEINHAUS, BASKIN, CLARK, & KIRBY

Figure 3. Representative absorbance - time spectral plot (typical example) showing lag phase for hemoglobin - NO_2^- reaction at 37°C and $\text{pH} = 7.4$. $[\text{Hb}] = 4.31 \times 10^{-5} \text{ M}$; $[\text{NO}_2^-] = 1.10 \times 10^{-3} \text{ M}$. Data taken every 0.2 min.

Figure 4. Plot showing dependence of pseudo-first-order rate constant upon NO_2^- concentration for reaction between myoglobin and NO_2^- at 37°C and $\text{pH} = 7.4$. Two NO_2^- concentration ranges are shown, the lower range in the inset.

Mb: The oxidation of Mb with nitrite ion, WR6026, and primaquine proceeded without precipitation, while using WR238,605 caused the formation of a precipitate. No kinetic data were obtained for the WR238,605 reaction. The reactions of Mb with WR6026, primaquine, and NO_2^- were all found to obey equation 5. The oxidation of Mb with NO_2^- was carried out using an NO_2^- concentration range from $1.9 \times 10^{-4} \text{ M}$ to $4.0 \times 10^{-3} \text{ M}$, which was always at least a 10-fold excess relative to the Mb concentration. Reactions at higher NO_2^- concentrations showed a lag phase prior to the more rapid oxidation reaction, similar to the spectral example shown in Figure 3. Again, only points after the first half-life were used. Excellent first-order plots of the data were obtained. As shown in Figure 4, a plot of the pseudo-first-order rate constant vs. NO_2^- concentration showed two definite linear regions, depending upon the NO_2^- concentration and allowed resolution of the two second-order rate constants listed in Table 1. A plot of the pseudo-first-order rate constant vs. low concentrations covering the range between $1.9 \times 10^{-4} \text{ M}$ and $8.0 \times 10^{-4} \text{ M}$, shown in the inset of Figure 4, was linear, thus showing first-order behavior in NO_2^- , and yielded a second-order rate constant of $0.179 \text{ M}^{-1}\text{s}^{-1}$. A plot of the pseudo-first-order rate constant vs. high NO_2^- concentrations, covering the range between $8.0 \times 10^{-4} \text{ M}$ and $4.0 \times 10^{-3} \text{ M}$, was also linear and indicated a different first-order rate pathway in nitrite ion with a second-order rate constant of $13.4 \text{ M}^{-1}\text{s}^{-1}$.

The oxidation of Mb with WR6026 showed first-order behavior in both Mb and WR6026. A plot of the pseudo-first-order rate constant vs. WR6026 concentration followed equation 7,

$$k_o = k_1 [\text{WR6026}] + k_2 \quad (7)$$

thereby suggesting a two-term rate law. Values of the rate constants obtained from the plot are listed in Table 1 and are $k_1 = k_{\text{WR6026}}^{\text{Mb}} = 5.1 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1}$ and $k_2 = k_{\text{Mb}} = 2.7 \times 10^{-5} \text{ s}^{-1}$.

The oxidation of Mb by primaquine showed kinetic behavior identical to that of WR6026. Rate constant values, listed in Table 1, are as follows: $k_1 = k_{\text{Mb}}^{\text{prima}} = 2.9 \times 10^{-2} \text{ M}^{-1}\text{s}^{-1}$ and $k_2 = 1.4 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$.

STEINHAUS, BASKIN, CLARK, & KIRBY

Cyano Complex Formation Kinetics: The reactions of cyanide ion with Mhb and with MMb were carried out in the presence of a 10-fold molar excess of oxidant. This allowed evaluation of any interference in the cyanide formation reactions due to the presence of the oxidant. All reactions followed equation 8,

$$\frac{d[\text{CN Metheme protein}]}{dt} = k_o[\text{Metheme protein}] \quad (8)$$

where k_o represents the pseudo-first-order rate constant and is given by equation 9.

$$k_o = k \frac{\text{CN}^-}{\text{heme protein}} [\text{CN}^-] \quad (9)$$

The reactions were run using 10-, 20-, and 30-fold molar excesses of cyanide ion. Excellent first-order plots were obtained in all cases, thus establishing first-order behavior in Mhb and MMb. Table 2 lists the data. Following equation 9, plots of the pseudo-first-order rate constant vs. cyanide ion concentration were all linear and passed through zero within experimental error, thus showing a single-term rate law that is first-order in cyanide. Table 3 lists the rate constants obtained. The cyanide ion concentration ranged from 8.6×10^{-4} M to 2.6×10^{-3} M for the Mhb reactions and from 3.8×10^{-4} M to 1.2×10^{-3} M for the MMb reactions.

Table 2

Cyanide Concentration Dependency of Observed First-order Rate Constant for the Reaction on Cyanide Ion with Mhb and MMb in the Presence of Oxidizing Agents

Mhb					
NO ₂ ⁻			WR6026		
[CN ⁻], M	pH	k _o , s ⁻¹	[CN ⁻], M	pH	k _o , s ⁻¹
8.60 x 10 ⁻⁴	7.47	0.421	8.60 x 10 ⁻⁴	7.44	0.412
1.72 x 10 ⁻³	7.42	0.803	1.72 x 10 ⁻³	7.43	1.05
2.58 x 10 ⁻³	7.41	1.21	2.58 x 10 ⁻³	7.43	1.21
Primaquine			WR238,605		
[CN ⁻], M	pH	k _o , s ⁻¹	[CN ⁻], M	pH	k _o , s ⁻¹
8.60 x 10 ⁻⁴	7.44	0.459	8.60 x 10 ⁻⁴	7.42	0.568
1.72 x 10 ⁻³	7.44	0.819	1.72 x 10 ⁻³	7.47	1.07
2.58 x 10 ⁻³	7.40	1.37	2.58 x 10 ⁻³	7.53	1.59

STEINHAUS, BASKIN, CLARK, & KIRBY

MMb

NO ₂ ⁻			WR6026		
[CN ⁻], M	pH	k _o , s ⁻¹	[CN ⁻], M	pH	k _o , s ⁻¹
3.83 x 10 ⁻⁴	7.52	0.198	3.83 x 10 ⁻⁴	7.59	0.196
7.65 x 10 ⁻⁴	7.57	0.431	7.65 x 10 ⁻⁴	7.52	0.447
1.15 x 10 ⁻³	7.57	0.608	1.15 x 10 ⁻³	7.55	0.617

Primaquine			WR238,605		
[CN ⁻], M	pH	k _o , s ⁻¹	[CN ⁻], M	pH	k _o , s ⁻¹
3.83 x 10 ⁻⁴	7.54	0.174	3.83 x 10 ⁻⁴	7.58	0.221
7.65 x 10 ⁻⁴	7.58	0.449	7.65 x 10 ⁻⁴	7.61	0.394
1.15 x 10 ⁻³	7.54	0.686	1.15 x 10 ⁻³	7.62	0.603

Table 3

Resolved Rate Constants for the Reaction of Cyanide Ion with MHb and MMb in the Presence of Oxidizing Agents. Std. Dev. are Given.

MHb		MMb	
Oxidizing Agent	k _{MHb} ^{CN}	Oxidizing Agent	k _{MMb} ^{CN}
NO ₂ ⁻	458 ± 8.4	NO ₂ ⁻	534 ± 43
WR6026	463 ± 160	WR6026	549 ± 62
Primaquine	529 ± 64	Primaquine	498 ± 26
WR238,605	549 ± 6	WR238,605	602 ± 68

DISCUSSION AND CONCLUSIONS

Concentration Effect Curves: The concentration effect curves involving the oxidation of Hb could only be obtained at a time near to the time of mixing due to a precipitate which formed later on during the course of the reaction. The curves show that early in the reaction, increasing the amount of oxidant increases the amount of MHb in a linear fashion. The amount of MHb produced between the time of mixing and the time of measurement, about 45 min., varies with respect to the oxidant used, but is small (less than 12%) in all cases. Sodium nitrite is the most efficient oxidant. For example, at an oxidant concentration equal to one-half of the total Hb concentration, sodium nitrite oxidized 1.7% of the Hb compared to 1.5% using WR6026 and only 0.4% using primaquine. The same order holds true at other concentrations of oxidant.

The results for Mb illustrate the same type of oxidative behavior early

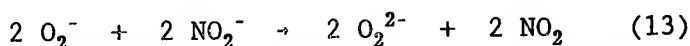
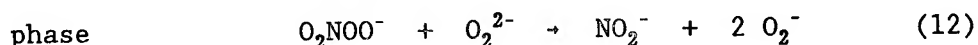
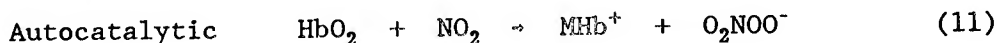
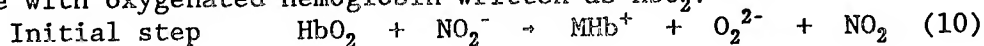
STEINHAUS, BASKIN, CLARK, & KIRBY

in the reaction, except that sodium nitrite is not the most efficient oxidant. For example, at an oxidant concentration equal to one-half of the total Mb concentration, NaNO_2 only oxidized 6% of the Mb compared to 11% using WR6026 and 12% using primaquine. At long reaction times, near equilibrium, and at oxidant concentrations less than 100% of the Mb concentration, the concentration effect curves show the formation of more Hb than expected. At 21 hours after mixing, solutions having 25%, 50%, and 75% oxidant relative to 100% Mb do not show formation of about 25%, 50%, or 75% MMb as would be expected for 1:1 stoichiometry but, rather, show formation of 80 - 90% MMb. Thus, either the stoichiometry is greater than 1:1 or there is another pathway for the formation of MMb. It is known that one mole of nitrite produces 2 equivalents of MHb from oxyhemoglobin¹⁹. Since there is only one heme group per molecule of myoglobin, one explanation for the data is that the reaction stoichiometry is 2:1 (Mb:nitrite). The stoichiometry of the reaction of the aminoquinolines with either Hb or Mb is not known; however, the concentration effect data obtained in this study suggest that it also is greater than 1:1, because at long reaction times 25 - 75% of either WR6026 or primaquine relative to 100% Mb produces 80 - 90% MMb just as nitrite does.

Oxidation Kinetics: Hb: The oxidation kinetics of Hb reacting with WR6026 and primaquine are obscured by the formation of a precipitate. The nature of this precipitate is unknown. Hb, MHb, and the oxidants used are all soluble in aqueous solution with the exception of WR238,605, which is soluble in DMSO. Upon initial mixing of the Hb and oxidant, no precipitate formed. About 1 to 2 hours after the reaction had been initiated and after the absorbance readings showed formation of MHb, the precipitate appeared. It is suggested that the precipitate is an insoluble complex between MHb and either the oxidized or reduced form of the oxidant. The formation reactions of MHb and cyanide appear to be independent of the particular oxidant present. This lack of interference could mean that the oxidant is not complexed at or near the heme sites where cyanide reacts. The rate of precipitation may follow the rate of MHb formation, meaning that as MHb is formed, precipitation occurs. The absorbance - time data obtained for these reactions is indicative only of the fact that the reaction rates depend on the oxidant concentration and that the reactions are slow. At a Hb concentration of 4.3×10^{-5} M and oxidizing agent concentrations in 10-to 50-fold molar excess, the reactions appear to last between 40 and 60 hours. This can be compared to the in vivo reactions of WR6026 and primaquine with Hb in beagle dogs, which shows a maximum formation of MHb at 4-7 days⁵. It is impossible to distinguish which oxidant reacts more rapidly. The reaction of Hb with nitrite proceeds without the formation of a precipitate. A lag phase, reported by other workers⁵⁻¹⁰, is seen at the beginning of the reactions. This lag phase has previously been reported to show a time dependence which is inversely proportional to the square of the nitrite ion concentration²⁰. The data obtained in this study show a direct relationship

STEINHAUS, BASKIN, CLARK, & KIRBY

between the lag phase time and the inverse of the nitrite ion concentration squared. Figure 3 shows a typical absorbance - time plot illustrating the lag phase followed by the rapid autocatalytic oxidation of Hb. As previously reported by others^{5,6,10}, our data showed the reaction to be first-order in both Hb and NO_2^- . Numerous studies of the NO_2^- oxidation of Hb have been reported⁵⁻¹¹. A fairly complete picture of the chemistry involved in the mechanism has emerged^{10,11}. Reactions 10 - 13 show the mechanism proposed by Doyle with oxygenated hemoglobin written as HbO_2 .



The initial or lag phase, reaction 10, involves a slow transformation of Hb to MHb which is first-order in both Hb and NO_2^- as found by others^{5,6,10}. The autocatalytic phase which follows, reactions 11 - 13, depends, among other things, upon catalysis by NO_2^- formed during the initial lag phase. The present study shows the oxidation autocatalysis also to be first-order in both Hb and NO_2^- . The above mechanism proposed by Doyle¹⁰ clearly shows first-order behavior in Hb, consistent with the experimental findings in this study. Since NO_2 is being continually generated by the first-order reaction of NO_2^- with Hb through the slow step which continues during the autocatalytic period, the autocatalytic phase must therefore be first-order in NO_2^- as seen in the present study. The present study thus confirms the proposed mechanism. Although no kinetic data for the WR6026 and primaquine reactions could be obtained for comparison with NO_2^- , it is obvious that the NO_2^- reaction proceeds much more rapidly.

Mb: The reactions of WR6026 and primaquine with Mb to produce MMb each proceeded through two pathways, one dependent upon and the other independent of the oxidant. The rate constants for the oxidant-dependent pathway, k_1 , of equation 7, show that primaquine reacts with Mb about 6 times faster than WR6026 does. The k_2 term of equation 7, which is independent of oxidizing agent, is also about 5 times larger for primaquine data than for WR6026 data. An explanation for this behavior is that the oxidizing agent acts as a catalyst for the self-oxidation of Mb, thus making the k_2 pathway independent of oxidizing agent concentration. However, primaquine may be a more effective catalyst, thus explaining the 5-fold difference in the k_2 values. The concentration effect curve data for Mb also supports this explanation. The formation of more MMb than expected, seen at long reaction times and attributed to higher than 1:1 stoichiometry, could also be due to a catalytic pathway that is independent of the oxidant concentration.

The oxidation of Mb by NaNO_2 appears to proceed through two pathways,

STEINHAUS, BASKIN, CLARK, & KIRBY

both dependent upon the NO_2^- concentration. The slower pathway found at lower NO_2^- concentrations does not show a lag period, whereas the faster pathway at higher NO_2^- concentrations does and behaves in a manner similar to the reaction of Hb with NO_2^- . It would appear that the lag period is necessary for acceleration of the rate of oxidation. The rate constant for the fast oxidation pathway is about 75 times larger than that for the slow pathway and is similar in magnitude to that for the oxidation of Hb by NO_2^- with Hb being oxidized about 7 times faster than Mb. The slow oxidation pathway is 6.5 times faster than oxidation by primaquine and 37 times faster than oxidation by WR6026. A plot of the pseudo-first-order rate constant vs. NO_2^- concentration for the slower pathway at low concentrations has an intercept statistically indistinguishable from zero, thus showing the absence of any catalytic pathway. Yet the dose response curve data at 21.25 hours shows the formation of 85% MMb if the curve is extrapolated to zero NO_2^- concentration, which suggests the presence of a catalytic pathway or a greater than 1:1 stoichiometry. These two facts are not reconcilable; no explanation for this behavior is available.

Cyanide Complex Formation Kinetics: Cyanide ion reacts rapidly with both MHb and MMb. Each iron atom in both heme proteins has only one available coordination site for the complexing of cyanide ion since five of the six coordination sites of the iron are strongly bound to heme and protein chain nitrogen atoms. Previous work shows that one mole of ferric heme is needed to bind one mole of cyanide ion²¹, thus making 4 the maximum value of x in equation 3. By analogy, Mb must react 1:1 with cyanide ion as shown in equation 4. The data in Table 3 shows the reaction mechanism to be insensitive to both the heme protein used and the presence of different oxidants. The similarity of the rate constants suggests three possibilities for the reaction mechanism. MMb contains only one heme unit while MHb has four; thus 1) only one heme iron in MHb may be reactive toward cyanide or 2) the rate-determining step occurs prior to the addition of the second cyanide ion to the second heme group. Further, at low concentrations, cyanide ion is known to be unreactive toward Hb²²; thus 3) the similarity of the MHb and MMb rate constants may suggest that only one Hb heme iron has been oxidized.

The mechanism of the formation of cyanomethemoglobin and cyanometmyoglobin should follow the Eigen-Tamm dissociative mechanism²³, which involves outer sphere association followed by a rate-limiting solvent dissociation with subsequent rapid coordination with the incoming ligand. The value of the formation rate for the reaction of NCS^- with Fe^{3+} , which follows this mechanism, has been found to be $1.3 \times 10^2 \text{ M}^{-1}\text{s}^{-1}$ ²⁴. The iron(III)- NCS^- reaction should be a good model for comparison with the cyanide reactions with MHb and MMb. It is likely that the positive charges of the heme sites not being attacked by the cyanide ion are too far removed to be a factor due to the large size of the protein such that the electrostatics of the reaction can be characterized as a +3, -1 interaction,

STEINHAUS, BASKIN, CLARK, & KIRBY

just as the Fe(III)-NCS^- reaction. The formation rate constants obtained in this study are all between 400 and 600 $\text{M}^{-1}\text{s}^{-1}$ and are quite similar to that reported for the $\text{Fe}^{3+} - \text{NCS}^-$ reaction. It is likely that the cyanide formation reactions also follow the Eigen-Tamm dissociative mechanism with rate of water loss as the rate-limiting step. Perhaps the most significant finding is that the MHb formed reacts rapidly with cyanide regardless of the oxidant. If the oxidants used in this study do bind to the metheme-proteins, they do so at a site sufficiently removed from the iron atom which binds cyanide or they are readily displaced by cyanide. Thus, the search for an effective antidote for cyanide may not be concerned with the effect that the antidote may have upon cyanide binding by the MHb sites.

Acknowledgements: The authors acknowledge the support of the USAMRICD, the U.S. Army Research Office, and the Battelle Research Triangle Park Office in carrying out this research.

REFERENCES

1. Kiese, M. (1974) Chapter 7 "Ferrihemoglobinemia Caused by Foreign Compounds." in Methemoglobinemia: A Comprehensive Treatise (Kiese, m., ed.), pp 55-168, CRC Press, Cleveland, Ohio.
2. Jandorf, B.J. and Bodansky, O. (1946) "Therapeutic and Prophylactic Effect of Methemoglobinemia in Inhalation Poisoning by Cyanide and Cyanogen Chloride." J. Ind. Hygiene and Toxicology 28, 125-132.
3. Chen, K.K. and Rose, C.L. (1952) "Nitrite and Thiosulfate in Cyanide Poisoning." J. Am. Med. Assoc. 149, 113.
4. Wittenberg, B.A., Wittenberg, J.B. and Caldwell, P.R.B. (1975) "Role of Myoglobin in the Oxygen Supply to Red Skeletal Muscle." J. Biol. Chem. 250, 9038-9043.
5. Wallace, W.J. and Caughey, W.S. (1975) "Mechanism for the Autooxidation of Hemoglobin by Phenols, Nitrite and 'Oxidant' Drugs." Biochem. Biophys. Res. Comm. 62, 561-567.
6. Radkey, F.L. (1976) "A Mechanism for the Conversion of Oxyhemoglobin to Methemoglobin by Nitrite." Clin. Chem. 22, 1986-1990.
7. Tomoda, A., Matsukawa, S. and Takeshita, M. (1977) "Effect of Inositol Hexaphosphate on Hemoglobin Oxidation by Nitrite and Ferricyanide." Biochem. Biophys. Res. Comm. 74, 1469-1474.

STEINHAUS, BASKIN, CLARK, & KIRBY

8. Kosaka, H., Imaizumi, K., Imai, K. and Tyuma, Y. (1979) "Stoichiometry of the Reaction of Oxyhemoglobin with Nitrite." Biochem. Biophys. Acta 581, 184-188.
9. Tomoda, A., Tsuji, A. and Yoneyama, Y. (1981) "Involvement of Superoxide Ion in the Reaction Mechanism of Haemoglobin Oxidation by Nitrite." 193, 169-179.
10. Doyle, M.P., Pickering, R.A., Dykstra, R.L., Nelson, C.L. and Boyer, R.F. (1982) "Involvement of Peroxide and Superoxide in the Oxidation of Hemoglobin by Nitrite." Biochem. Biophys. Res. Comm. 105, 127-132.
11. Doyle, M.P., Herman, J.G. and Dykstra, R.L. (1985) "Autocatalytic Oxidation of Hemoglobin Induced by Nitrite: Activation and Chemical Inhibition." J. Free Radicals In Biology and Medicine 1, 145-153.
12. Anders, J.C., Chung, H. and Theoharides, A.D. (1988) "Methemoglobin Formation Resulting from Administration of Candidate 8-Aminoquinoline Antiparasitic Drugs in the Dog." Fund. App. Toxicology 10, 270-275.
13. Smith, R.P. (1967) "The Nitrite Methemoglobin Complex - Its Significance in Methemoglobin Analysis and its Possible Role in Methemoglobinemia." Biochem. Pharm. 16, 1655-1664.
14. Van Assendelft, O.W. and Iljstra, W.G. (1965) "The Formation of Haemoglobin using Nitrites." Clinica Chim. Acta 11, 571.
15. Van Assendelft, O.W. (1965) "The Reaction of Myohaemoglobin and Nitrite." Clin. Chim. Acta 12, 546.
16. Dershwitz, M. and Novak, R.F. (1982) "Generation of Superoxide via the Interaction of Nitrofurantoin with Oxyhemoglobin." J. Biol. Chem. 257, 75-79.
17. Eilers, R.J. (1967) "Notification of the Final Adoption of an International Method and Standard Solution for Hemoglobometry Specifications for Preparation of Standard Solutions." The Am. J. of Clin. Path. 47, 212-214.
18. Benesch, R.E., Benesch, R. and Yang, S. (1973) "Equations for the Spectrophotometric Analysis of Hemoglobin Mixtures." Anal. Biochem. 55, 245-248.

STEINHAUS, BASKIN, CLARK, & KIRBY

19. Austin, J.H. and Drabkin, D.L. (1935) "Spectrophotometric Studies III. Methemoglobin." J. Biol. Chem. 112, 67-88.
20. Kakizaki, T., Sato, M. and Haseyawa, H. (1964) "Oxidation of Oxyhemoglobin by Sodium Nitrite." Ind. Health 2, 124-126.
21. Ref. 1, p. 62.
22. Smith, R.P. and Cosselin, R.E. (1966) "On the Mechanism of Sulfide Inactivation by Methemoglobin." Toxicology App. Pharm. 8, 159-172.
23. Ref. 1, Chapter 2, p. 3.
24. Margerum, D.W., Cayley, C.W., Weatherburn, D.C. and Pagenkopf, G.K. (1978) Chapter 1 "Kinetics and Mechanisms of Complex Formation and Ligand Exchange." in Coordination Chemistry (Martell, A.E., ed) pp 7-18, American Chemical Society, Washington, D.C.

Military Robotics: A Technology Assessment From a User's Perspective (U)

Major Mark L. Swinson, Ph.D., P.E.
HQ TRADOC
Fort Monroe, Virginia

1.0 INTRODUCTION

Reductions in manpower, both announced and anticipated, do not bode well for our stated intent to improve (or at least roughly maintain) American conventional military capability. This issue of conventional forces was recently highlighted both because of a treaty mandated, reduced reliance upon theater nuclear weapons, as well as an increased emphasis with regard to third world contingencies. Additional manpower reductions will almost certainly exacerbate the difficulties already experienced by planners in their efforts to restructure our Army, while at the same time seeking to maintain a balanced mix of capable, deployable and sustainable forces. Some experts believe that robotics may offer a partial solution.

Very human-like (anthropomorphic) robots certainly have been a recurring theme of science fiction writers for as long as most can remember, and will no doubt continue to be so for many years to come. However, as any auto worker knows first hand, robots are not just the stuff of fantasy, they're here today.

Of course, the science of robotics encompasses more than just the welding or spray painting of auto bodies, else it would be of little interest to military planners. Instead, the term "robotics" is generally taken to mean the broad range of technologies which find one of their principal (or at least one of their more exciting) applications in "robotic systems." These technologies generally include manipulators, actuators, digital controllers, sensors, data links and "artificial intelligence."

Taken together, these technologies portend the capability to perform certain necessary military missions with reduced manpower requirements. These missions might include those which are especially manpower intensive, such as certain logistical operations; those involving extremely hazardous tasks, such as explosive ordnance disposal (EOD); or those missions performed in especially hazardous environments, such as

those tasks performed in space, under the ocean, in hostile skies, or in the beaten zone of a modern battlefield.

Unfortunately, the central problem with military robotics, like most of the other so-called "emerging technologies," is that they never seem to emerge. As long as that's true, no field experience is gained. Without field experience, neither insight nor constituency from the user is likely to develop. Consequently, definitive user requirements are not forthcoming.

The most effective way to correct this problem is to rapidly field a relatively simple, modestly capable, comparatively low cost system in significant numbers. That is, exploit available technology! Consequently, this assessment focuses primarily on currently available robotics technologies which could support the rapid fielding of military systems.

It should be noted that, strictly speaking, most "military robots" aren't "robots" at all, but instead are teleoperated (remotely operated) vehicles. These military systems basically come in one of three varieties: unmanned aerial vehicles (UAVs), unmanned underwater vehicles (UUVs), and unmanned ground vehicles (UGVs), and may include both robots and teleoperators. The following definitions were used for this assessment.

- Teleoperator. A mechanical system, mobile or stationary, that requires a human in the control loop on a real-time basis to remotely perform part or all of the task.

- Robot. A mechanical system that performs a task involving sensing, computation and action independently of real-time human control.

Military robots, therefore, are those unmanned systems (or subsystems) which permit the soldier to continue to perform his mission, though separated from the job site, in time or space. As such both teleoperators and robots will be considered as "military robotic systems" for the purpose of this analysis.

2.0 SURVEY OF CONSTITUENT TECHNOLOGIES

2.1 MANIPULATORS

For most people, the term robot conjures up the image of an anthropomorphic manipulator. In fact, the anthropomorphic bias is clearly prevalent in nearly all of our robotics concepts. That is, we typically envision automating a task currently performed by humans with no fundamental change in the nature of the task, itself. This is perfectly understandable, of course, since most of our exposure to actual robots comes from the wide application of industrial robots in the

automobile industry. In the highly engineered environment of a modern automobile manufacturing plant, industrial robots readily perform boring, fatiguing and dangerous work formerly performed by a highly unionized, blue-collar labor force, with only the most superficial changes in the essential nature of the production line. However, with the exception of a few applications in the depots, tank plants, ammunition plants or other industrial type operations, unmodified industrial type manipulators are probably of little direct utility for military applications.

Nonetheless, it would be incorrect to assume that this particular "robot incarnation" was the state-of-the-art for manipulators, or even the only type of commercially available manipulators! For example, two of the most widely promulgated metrics for manipulators are the payload-to-weight ratio and the payload capacity. Both have typically been small for first generation, industrial robots, due to the need for structural rigidity. Most industrial robots use collocated servo sensor systems for closed loop control. This means that the position sensor is located at the actuator, rather than on the end effector. Consequently, the location of the end effector must be inferred from joint position data, hence the need for structural rigidity. Such simple designs enhance stability and lower production cost, but also limit the utility and range of potential applications.

Recent advances in design, however, have dramatically improved performance on both counts. For example, a manipulator with a one-to-one payload to weight ratio, built by Odetics of California, was delivered to the Human Engineering Laboratory (HEL) in 1989 and incorporated into the Soldier Robot Interface Project (SRIP) test vehicle. See figure 1. The Field Materiel Handling Robot (FMR), built by Martin Marietta for HEL, is a heavy lift system designed for handling pallets of artillery ammunition. See figure 2. While still inferior to biological systems in many ways, such manipulators are adequate for a wide range of demanding applications.

Another issue is the degree of autonomy associated with the manipulator system. It is worth pointing out that teleoperated manipulators, or remotely operated arms, are more likely to provide immediate military application than fully preprogrammed systems. The reason is that few military applications provide the kind of structured environment necessary for totally unattended, playback-type operation. While research into sensors, sensor fusion, and artificial intelligence offers some interesting possibilities, this research is not likely to soon provide sufficiently sophisticated products in order to provide the option of fully autonomous operation, at least in an unstructured, tactical environment.

It is perhaps worth noting that the initial development of manipulators occurred as part of the nuclear energy program. These early mechanisms were used to permit an operator to remotely handle highly

radioactive materials. Because the human operator is making all the high level control decisions, that is, because he is in real-time control of the manipulator, he retains the ability to operate in an uncertain and/or unstructured environment. The payoff comes from removing the operator from a dangerous environment, as opposed to removing the operator from real-time control of the process.

Such teleoperated devices are currently used in a wide range of hazardous environments. Examples include explosive ordnance disposal, firefighting, undersea maintenance and servicing, space shuttle operations and operations in a contaminated environment. Many of these tasks are also characteristic of military operations. The near-term payoff for teleoperated manipulators comes from removing the operator from the point of danger and the consequent reduction in the number of casualties.

Teleoperated manipulators are extremely cost effective since their application to tasks performed in hazardous environments minimizes the need for costly, cumbersome, and sometimes tragically ineffective protective clothing. While this category of tasks may be somewhat limited, there are a considerable number of opportunities yet to be exploited by the Army. Furthermore, the experience (and confidence) gained from the widespread use of such equipment will facilitate the integration of more capable manipulators. Fairly near-term systems will almost certainly permit a few personnel to control many such machines performing what previously had been very labor-intensive tasks.

2.2 ACTUATORS

An essential issue in the design and realization of any robotic system which incorporates articulation is the choice of actuator design. There are three general types, namely pneumatic, hydraulic and electric.

Pneumatic actuators are operated by air pressure, and generally have a two position arrangement. Hence, they are often referred to as "bang-bang" actuators. More exotic designs, such as bladder arrangements, where air pressure corresponds to the extent of manipulator erection, have been used for selected applications, such as automotive paint spraying. The main advantage of pneumatic actuators is that they are clean (no leaky hydraulics) and safe for flammable or explosive environments (no sparks from electric motors).

Hydraulic actuators typically use some sort of piston/cylinder arrangement with a liquid working fluid, usually oil. Such actuation has long been used in the construction engineering and manufacturing world, since hydraulic systems offer high power densities. In fact, operating at pressures of 4500 psi, there is no power transmission medium denser than fluid power. The main drawbacks are leakage and relatively poor positional resolution.

Recent developments have largely overcome these deficiencies. Better

design has dramatically reduced the leakage problem, while the interface of digital electronics has improved articulation performance. In fact, the combination of digital controls and closed-loop technology has made it possible to achieve the levels of accuracy and repeatability required (on the order of 0.0005 inch) for many manufacturing process applications. The significance of this so-called "electrohydraulic motion control" is that it permits the automation of some large scale tasks beyond the capacity of current electric motors. The automation of a John Deere excavator as part of the Air Force's Rapid Runway Repair (RRR) Project is one example. See figure 3.

Nonetheless, electric motors remain the most important actuation means for advanced applications, and especially for future developments. Electric motors are well suited to digital control, and through the use of well designed transmissions (such as harmonic drives) can provide excellent performance. Further, electric motor performance is continually improving.

Because of the broad range of commercial applications, advanced actuators are readily available. The result is that new opportunities for task automation are continually facilitated, particularly for developing new (though not necessarily anthropomorphic) manipulator systems. What's more, special purpose automation also has become more cost effective, since little or no developmental efforts are required for the actuator subsystem. Autoloaders are one military application.

Another is a current developmental effort underway at the Belvoir RD&E Center called the Universal Self-Deployable Cargo Handler (USDCH), which is an example of the possible use of these actuators to articulate a fairly sophisticated mechanism. See figure 4. Specifically, part of this forklift vehicle design calls for an automatic pallet acquisition system which makes use of an array of computer-controlled actuators. Since pallet acquisition is a fairly difficult task, automating this operation reduces the required training level for the forklift operator, as well as potentially speeding system operation.

2.3 DIGITAL CONTROLLERS

Digital controllers are one of the most exciting and rapidly changing robotics technologies. These controllers generally come in two varieties, microcomputers and micro-controllers.

Microcomputers are usually board-level designs, with each board capable of providing programmable servo-control for one to three actuators. Micro-controllers, on the other hand, are single chip microcomputers, such as the Motorola 68HC11 and the Intel 8096, with the capability to provide single axis control. These devices permit real-time control and typically include various I/O (input/output) options, expansion devices and memory expansions. Interrupt timing constraints are a critical design parameter since real-time, digital

control is characteristically an intensely interrupt-driven environment.

The key points about digital controllers are that they facilitate the automation of equipment that previously had to be totally manually operated, as well as reduce the work load of the human operator. This is true because the controller can transparently handle the low level servo-control tasks and thus free the human operator to concentrate more on the high level decision making.

One example is the previously mentioned RRR Project at Tyndall Air Force Base, developed in conjunction with the University of Florida's Center for Intelligent Machines and Robotics (CIMAR). The focus of this effort to date has been the automation of a John Deere excavator for the repair of bomb damaged runways. This contingency mission would likely have to be performed in a contaminated, ordnance strewn environment. Automating the operation of the excavator frees the operator to concentrate on task planning without having to physically manipulate the manual controls in the vehicle cab (something that would be very difficult in full protective ensemble). Additionally, operator training time is reduced since the actual excavator arm operates under computer control. Think of it as an automatic transmission. The operator is still "driving the vehicle," he just doesn't have to worry about "manipulating a stick shift or a clutch."

This is an important point. By automating such tasks as digging, compacting, tool changing, etc., one not only can remove the operator from the cab (very desirable for working on a contaminated, ordnance strewn runway) but also reduce the level of training required of the operator. If done well, the more sophisticated the design, the more transparent its application for the user. Thus, good design should reduce, not raise, operating and sustainment (O&S) costs.

Such digital controllers are readily available, and very reasonably priced. Because of their wide application in the commercial sector, there is little or no requirement for the expenditure of scarce R&D funds by Army agencies for development.

2.4 SENSORS

The most militarily significant development in sensors in recent years may well be the development of relatively low-cost television cameras. When combined with a broad bandwidth data link, such as fiber-optic cable, the available high quality, real-time video feedback offers tremendous opportunities for remote viewing and operations. The operator need no longer have direct observation of a system to be controlled or task to be performed, but instead can rely upon video feedback. The Fiber Optic Guided Missile (FOG-M) is an excellent example.

At the heart of these cameras is what is called a charge-coupled device (CCD), which is nothing more than an integrated circuit chip that

can store localized "packets" of charges, and then transfer them to adjacent locations as required. The output charges can represent digital information. Current systems can operate in the 5-7 MHz range.

However, the process of image interpretation is one that is more suited to humans than machines. Image interpretation, once the darling of the artificial intelligence community, is still largely a phenomenon of the laboratory, not the factory floor. Hence, the main role of the camera is to facilitate "telepresence."

Telepresence refers to the projection of operator perception predicated on real-time, sensor feedback provided to a remote operator. While current developmental efforts in the area of human factors engineering continue to examine the benefits of various sensory cues, such as force/torque and acoustic feedback, video feedback remains the quintessential characteristic of telepresence, and hence the most important technology opportunity for the near-term fielding of a tactical unmanned ground vehicle. The Teleoperated Mobile All-Purpose Platform (TMAP) is representative of such a vehicle. Two versions of the TMAP were built for MICOM, one by Martin Marietta and one by Grumman. See figure 5.

Teleoperated vehicles are often confused with robots, and for good reason. Many of the same technologies are incorporated, especially advanced actuation, sensors and computer control. However, as was stated in the definitions, the fundamental difference lies in the nature of operator involvement in the control loop. Teleoperated systems require real-time human control; robots, which are preprogrammed, do not. It should also be noted that there is arguably a class of systems somewhere in between these two which is usually referred to as telesupervised.

Telesupervised systems normally operate autonomously (preprogrammed) but are designed to easily accommodate real-time operator intervention, especially for high level decisions. Given the current level of work in computer control, telesupervision represents a relatively near-term evolution of telepresence technology. One application could be the "follow-the-leader" concept. The idea is to have a column of vehicles automatically follow the path of a manned lead vehicle with the required navigational information provided to the following vehicles via some type of digital data link. Such a system would require some additional autonomous capability for dynamic obstacle avoidance, as well as the facility for emergency manual override. The Training Wheels program developed by Honeywell for PM TRADE envisions such a system. This effort is to simulate second echelon OPFOR Regiments at the National Training Center without a corresponding increase in personnel to support the OPFOR.

Besides vision, there are other sensors which offer significant opportunities. Acoustic sensors are one example. Unlike vision, acoustic sensors offer an excellent opportunity for early exploitation of autonomous system capabilities.

Military vehicles, especially helicopters, generate quite distinct signatures that are difficult to alter or mask. Further, a comparatively small amount of data is generated by such sensors as compared to video feedback. Consequently, acoustic sensors require only modest, on-board computational capability. Ongoing development efforts in the area of wide area mines (WAMs) are indicative of efforts to capitalize on this opportunity.

2.5 DATA LINKS

In order to facilitate real-time human control a bi-directional data link is needed. Bandwidth is the key issue, especially for video feedback, since continuous video signals require a much wider bandwidth than is available using tactical radio frequency (RF) data links. Two alternative approaches are currently being pursued.

One is "data compression." This term is used to refer to techniques which modify the quantity of data required to be transmitted, either via hardware or software, so as to permit the use of conventional, RF data links. The main advantage is the (more or less) non-line-of-sight capability of tactical radios. The main disadvantages are the quality of the video feedback, signal vulnerability, and the lack of frequencies available for use in this range of the electromagnetic spectrum.

With imaging sensor systems, the easiest way to achieve data rate reduction is by reducing the transmitted frame rate. The limitation of this approach is that to perform closed loop operations in real-time, a frame rate of about 8 frames per second is the minimum acceptable. This still exceeds the capability of tactical RF systems.

To further reduce the required bandwidth, the image can be processed such that redundant scene information is not transmitted. Of course, different algorithms work in different ways. The key is to transmit the minimum number of bits per second, while at the same time minimizing image degradation.

The alternative to data compression is to use a broadband data link. Current systems rely primarily upon fiber-optic cable, though research work in the U.K. using laser and microwave continues. Of some importance is the fact that none of these broadband alternatives are heavily subscribed by current tactical systems. The main drawback of laser and microwave is the need for line of sight, though repeaters (whether ground based or airborne) can overcome this limitation. The other problem is penetration, though some newer lasers penetrate fog and dust as well or better than tactical RF signals can.

Fiber-optic cable for these systems is different than that used for telephone lines, primarily in the ruggedness of design. Currently available fiber-optic cable systems provide a crisp, clear image that is both secure and electromagnetic pulse (EMP) hardened. The bandwidth available over fiber-optic cable is more than sufficient to carry stereo

video signals, as well as providing additional bandwidth to accommodate additional future requirements. The main drawback is the physical tether, though this limitation may be more psychological than physical.

2.6 ARTIFICIAL INTELLIGENCE

This term has been so often associated with robotics, at least within DoD, that no attempt at a comprehensive discussion of robotics technologies dare ignore it. While a widely accepted definition for the term "artificial intelligence" (AI) is considered by many to be emotive, I refer here to that body of knowledge about "computing" and "intelligence" which purports to obviate the need for a man in the loop to make even the "high level decisions" (such as shoot or don't shoot).

Ever since the development of "expert systems" in the early 1980's, AI has been highly touted as one of those new technologies that would revolutionize nearly everything we do. The "unmanned factory" was thought to be only a few years away. Human involvement in most tasks seemed to be on the verge of redundancy. Unfortunately, as noted by Karl Kempf at the Eleventh International Joint Conference on Artificial Intelligence, about half the money spent on AI so far has produced nothing more than a collection of "very expensive conference papers." He suggested that much of this failure was due to the "misapplication" of AI.

Kempf suggested that there were three classes of AI applications. The first was simply the codification of reasoning for tasks human experts understand well, thus allowing novices to perform like experts. Failure diagnostics for automotive repair is an example. The second class involves those tasks which are also well understood, but are difficult for humans to solve due to the large volume of data to be manipulated, or the rapidity with which the data arrives to be manipulated. HQ TRADOC recently used such a system to help in the compilation and formulation of the Long Range Research, Development, and Acquisition Plan (LRRDAP). Finally, the third class consists of the most glamorous problems, namely those which we humans really don't know how to solve very well. Those applications are almost certainly doomed to failure. It is this last type of problem, such as controlling fully autonomous robots, particularly in relatively unstructured environments, which has been oversold.

3.0 ANALYSIS

While fully "autonomous" systems garner most of the glamour associated with robotics, the battlefield is an environment that offers relatively few opportunities for their near-term application. It is simply not possible to sufficiently engineer the battlefield environment to accommodate the limited capacity of contemporary machine intelligence

for fully autonomous operations. However, the situation for teleoperators is another matter altogether.

Unfortunately, the ready availability of the basic technology to build simple teleoperated systems has led some to incorrectly conclude that these systems are therefore somehow trivial, and certainly inferior to "real robots." That conclusion is incorrect. On the contrary, teleoperated systems offer dramatic, high-payoff, but unfortunately as yet unrealized opportunities, not so much for lack of technology as for lack of constituency, and especially for lack of a champion.

To the uninitiated, there is a common misperception that if one can see the task, one can do it. The fact is, meaningful tasks in the real world are always darn hard! Teleoperation is NOT a done deal. Real challenges remain, but potential payoffs are enormous.

The commercial use of teleoperated systems is rapidly growing and consequently, so too is commercial availability. The Army has recently purchased a number of explosive ordnance disposal (EOD) robots of the type which are readily available to law enforcement and security agencies. See figure 6. Similarly much, if not most, of the undersea maintenance and exploration done today is conducted by UUVs. NASA has also focused research efforts in this arena, such as the Flight Telerobotic Servicer (FTS), due to the cost and inherent limitations of having astronauts perform extra vehicular activities (EVA).

Clearly, teleoperators offer a tremendous opportunity to perform selected operations in highly dangerous environments with little or no danger to the operator. (It is hard to imagine a more dangerous environment than the beaten zone of an NBC contaminated battlefield.) Furthermore, these systems clearly have the capability to grow in sophistication and range of application as more and more of the low level control functions are automated, and as user confidence grows to the point where tasks are redefined to more readily accommodate machines.

I do not want to suggest, however, that teleoperation is a panacea for solving our defense needs, any more than nuclear weapons have been. Rather, teleoperated systems offer a range of unique capabilities and unique developmental challenges. Simply stated, teleoperated systems can do real things, today, not just pedagogically interesting demonstrations in the laboratory.

Conventional wisdom suggests that reconnaissance, surveillance and target acquisition should be the first tactical mission assigned to UGVs. There is precedent. As we all know, this role was the first tactical mission assigned to manned military aircraft, as well as the current generation of unmanned aerial vehicles (UAVs).

Obviously, the capability to directly view the battlefield via video feedback is exciting. Perhaps not since before the American Civil War has the battlefield commander had the opportunity to directly observe areas of interest, conduct route reconnaissance, and direct fire support without having to continually expose himself or his subordinates to

direct fire.

This opportunity largely derives from the capability to safely relocate an array of sensor suites on the battlefield on a real-time basis. This capability, combined with the ongoing efforts in sensor fusion, portends the opportunity for a commander to have a continuous, accurate, multidimensional, real-time "picture" of the battle.

Unmanned vehicles may also play a significant role in weapons employment. The signatures produced by modern weapons are such that the weapons platform is very susceptible to counterfire. Similarly, laser designation is a very dangerous task for a manned platform. Both of these missions were successfully demonstrated by unmanned ground vehicles at the Camp Pendleton UGV "advanced technology demonstration" in September 1989.

Another issue which is usually ignored during peace time is that of fratricide. This issue has been a major stumbling block to the employment of certain advanced weapons concepts, such as high powered microwave (HPM). Indeed, even the initial reconnaissance module (which includes video and acoustic feedback) could reduce the danger of tragic mishaps during particularly dangerous maneuvers, such as link-up operations and passage of lines during periods of reduced visibility. Also, by providing the capability to permit unidentified objects to sufficiently close with an unmanned reconnaissance platform for positive video identification without risking the lives of US military personnel, unfortunate accidents (like the shooting down of the Iranian Airbus mistaken for an Iranian warplane) might also be avoided.

A fairly large number of other mission packages are also under consideration, including NBC reconnaissance, intrusion detection, countermine, smoke generation and material handling. The point is simply that unmanned platforms can dramatically alter the cost/benefit ratio used to analyze the feasibility of various tactical, as well as doctrinal options.

4.0 RECOMMENDATIONS

There is little doubt that "robotics," like most new technologies, has been consistently oversold. For at least ten years, robotics has been "just around the corner." Consequently, as one might well expect, interest in the user community has begun to wane, as has support. This problem is aggravated by the current environment of rapidly diminishing resources.

This situation is not only unfortunate, but arguably even dangerous. As we have seen in the UAV arena, our allies' efforts have far exceeded our own, as have their accomplishments. Fortunately, our current joint program for UAVs appears to be on the right track. Hopefully we can learn from our past mistakes, and do better in the UGV arena.

Technology opportunities clearly exist. The most important challenge

seems to be the formulation and execution of a coordinated exploitation strategy under centralized leadership. This individual would serve as the champion so desperately needed to give the fledgling robotic systems a chance for programmatic survival. While some form of centralized management appears to have been mandated by Congress, the final form has yet to be fully realized. However, the temptation to vest a committee with this responsibility should be steadfastly avoided. Such committees are generally perceived as little more than attempts to diffuse the consequences of failure.

A major goal of the management team must be to focus our limited resources. The PEO/PM should consider the following four-part approach. The first (and most important) is to strongly support the tactical UGV effort. The goal here is to rapidly field a teleoperated unmanned ground vehicle for tactical units. This program is essential. Obviously, such platforms offer a long endurance "look over the next hill" capability so desperately needed by our maneuver forces. Perhaps even more important, however, is the operational experience (and hopefully broad-based constituency) to be gained from the tactical employment of UGVs. Consequently, this program will help ensure the efficacy of our long-term developmental efforts.

Secondly, a comparable effort should be made to develop logistics oriented systems so as to achieve similar kinds of objectives. Note that a similar recommendation was also made by the National Research Council 1987 Study entitled Army Robotics and Artificial Intelligence. This logistics oriented effort should include subsystem development and encompass both those systems that facilitate a reduction in O&S costs, as well as those that enhance sustainment capabilities, particularly in support of the more demanding environment of a non-linear battlefield.

The third part is to continue our Non-Developmental Item/Commercial Off-The-Shelf (NDI/COTS) type procurement efforts (such as for EOD). Such systems are commercially available, and should continue to be acquired for Army use despite the objection of "not invented here."

The last part of this program is to develop and manage a centrally coordinated, responsive research effort in order to ensure the maximum possible return for our scarce R&D dollars. This will hopefully preclude much of the duplication and confusion that has been characteristic of our past efforts.

It should be noted that three of the four efforts suggested are oriented toward initial systems fielding rather than research. The sad fact of the matter is that we have not done well at capitalizing upon currently available technology. In fairness, it should be noted that our civilian industry has demonstrated a similar failing, with our degraded market share in several industries, most notably in consumer electronics and automobiles, being the direct result. Unfortunately, when the subject is defense technology, the potential consequences of failure are much more than simply lost market share.

SWINSON

The ultimate goal of the PEO/PM then should be to facilitate the development of a wide base of user support, based upon the rapid, successful fielding of initial systems, a viable ongoing developmental effort, both for tactical and logistical systems, and a rational, coordinated research effort to ensure technological viability of the overall effort. Such a plan seems to offer the highest probability that the Army will be able to capitalize upon this emerging technology which will almost certainly impact upon the fundamental nature of land warfare.

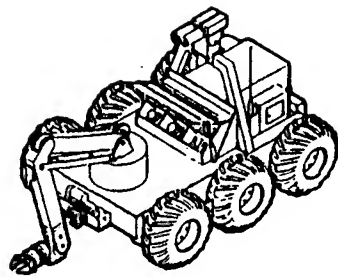


Figure 1 SRIP

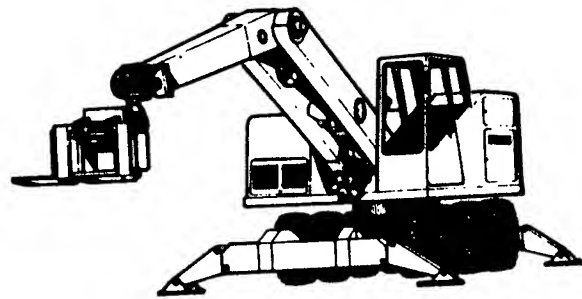


Figure 2 FMR

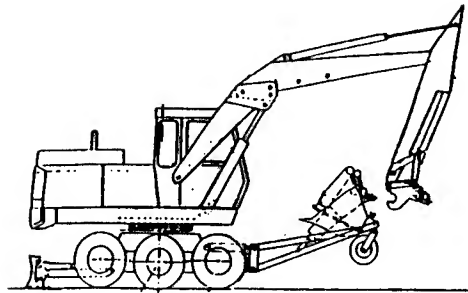


Figure 3 RRR

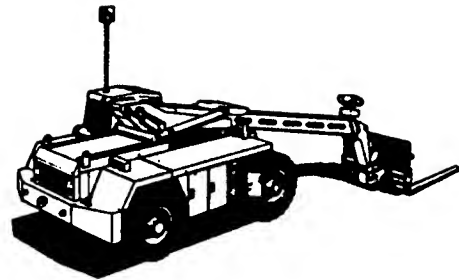
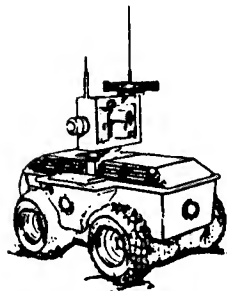
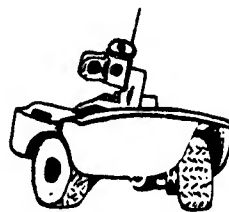


Figure 4 USDCH



a)



b)

Figure 5 TMAP
a) Martin Marietta
b) Grumman

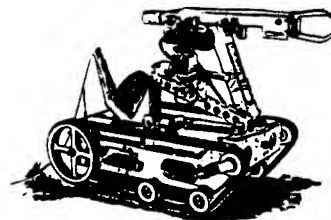


Figure 6 EOD

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

An Innovative Fluid Delivery System for Chemical Protective
Clothing: An Evaluation (U)

Patricia C. Szlyk*, Dr., Ingrid V. Sils, Ms., Ralph P.
Francesconi, Dr., Roger W. Hubbard, Dr.
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760-5007

During training exercises or maneuvers, soldiers will not always consume sufficient water to compensate for the fluid lost in sweat (1,12,15), despite ample supplies of water. This results in voluntary dehydration. The level of dehydration is influenced by many factors including work load, environmental temperature, water accessibility and palatability (temperature and taste), and adequate time for drinking (1,5,12,13). Ordinarily, the more intense the voluntary dehydration, the greater the decrement in a soldier's performance in the field. Hypohydration levels ranging from 1.5% to 7% of initial body weight have been observed in both laboratory and field settings (1,11,12) during work under hot environmental conditions, even though water was plentiful, palatable and readily available.

Under the threat of biological, chemical or nuclear contamination, the soldier must function while encapsulated in chemical protective gear and hooded M17A1 mask. The addition of this clothing inherently increases sweat rates, hypohydration and hyperthermia, even in temperate conditions (3,14,18). An additional constraint placed on the soldier is the difficulty of drinking with the current system.

Utilization of the current through-mask drinking system (CS) has been described as lengthy, cumbersome, frustrating, and generally difficult because it requires two-handed operation, decontamination and connection/disconnection with each drink, and first creating and then drawing against a positive pressure. Thus, an innovative hand-pump driven drinking system (FH) was developed to lessen some of the risks and problems associated with the current through-mask system. This fluid delivery system requires a single connection to the canteen and mask, and employs a single-handed squeeze bulb that hydraulically moves water from the canteen to the drinking spigot of the mask.

The current study was conducted to compare fluid balance, heat

UNCLASSIFIED

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

strain and subjective ratings of soldiers working in chemical protective gear while using either the current drinking system or the hand-pump driven system.

METHODS

Eighteen unacclimated males served as test subjects for this evaluation. All subjects were briefed on the purpose and risks of this study, underwent a thorough medical screening, and provided written consent prior to participation.

Subjects were randomly assigned to one of two experimental groups, differing only in the fluid delivery system used for drinking. One group was instructed to drink using the current through-mask system (CS) while the second group consumed fluids using the hand-pump driven fluid delivery system (FH).

On the test day, subjects reported to the test facility at 0645h where they consumed a commercially available instant breakfast in 450 ml milk, toast with butter or jam, and 450 ml of orange or apple juice. Subjects then proceeded to a dressing room where measurements of initial nude body weight, age and height were taken. Physical characteristics were similar for both groups of subjects (Table 1). A pre-test urine sample was analyzed for specific gravity (refractometry) in order to assess hydration level (1.024 ± 0.001). Individuals with a urine specific gravity ≥ 1.030 were instructed to drink additional fluids prior to entering the chamber. After placement of instrumentation for measuring rectal (Tre) and skin (Tsk) temperatures as well as ECG electrodes to measure heart rate (HR), subjects donned a T-shirt, shorts, socks, and sneakers. Over this clothing chemical protective jacket, trousers, boots, gloves, and hooded M17A1 mask (MOPP IV configuration) were worn.

Drinking water (25°C) was iodinated (16 mg I_2 /L) to simulate field pick-up water and was supplied in 2 qt collapsible canteens that were carried on the waist attached to web gear. Fluid intake was determined by weighing canteens at the beginning and end of each work and rest cycle on an electronic balance (± 1 g). In order to compare the fluid intake data, all fluid intakes were normalized to L/hr to account for the differences in duration of the walk and rest cycles, as well as endurance times. Sweat loss was calculated as the difference in nude body weight from pre- to post-exercise, adjusting for fluid intake and urine output.

Subjects walked (4 km/h, level treadmill) for six consecutive 50/10 minute work/rest cycles in a warm test chamber (32.6°C dry bulb, 17.5°C wet bulb, 20% relative humidity, and 8.05 kph windspeed, producing a WBGT of 22.1°C) while carrying a rifle.

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

Table 1. Physical characteristics of subjects

Variable	CS (n=9)	FH (n=9)
Age (yrs)	23.6 \pm 3.1	28.0 \pm 2.1
Height (cm)	178.2 \pm 1.7	180.0 \pm 2.4
Weight (kg)	78.8 \pm 2.6	76.0 \pm 3.0
Surface area (m ²) ¹	1.95 \pm 0.04	1.98 \pm 0.04
Body fat (%) ²	16.5 \pm 2.0	18.6 \pm 1.4

Values are Mean \pm S.E.

1 Calculated from height and weight measurements (7).

2 Durnin-Womersley (8) formula for percent body fat, using four skinfold sites and adjusting for age.

CS Current through-mask drinking system.

FH Hand-pump driven system.

Subjects were sedentary during the remaining 10 min of each h. Subjects carried a rifle during walks and prior to drinking with the CS, soldiers slung the rifle on their shoulders. CS subjects were instructed to decontaminate mask and canteen connections, then connect and disconnect couplings with each drink. FH subjects decontaminated connections only once during the trial because the mask drinking tube and canteen remain attached to the hand-pump tubing with this system.

All subjects were asked to complete an Environmental Symptoms Questionnaire (ESQ) while in MOPP IV configuration before starting their first chamber walk and immediately following their final walk. The ESQ consisted of a series of questions which were rated on a 6-point scale from "not at all" to "extreme" and reflected each subject's perception of his condition during the heat/exercise scenario. Subjects were also asked to evaluate the ease of use of their drinking system during the 2nd, 4th and 6th or final rest period. Perceived exertion was measured at the end of each exercise cycle; each subject described the difficulty of his walk in increments ranging from "very, very light" to "very, very hard" (4).

RESULTS AND DISCUSSION

The hand-pump driven water delivery system was developed to alleviate some of the limitations to drinking imposed by wearing chemical protective clothing and the hooded mask, and using the current system. Several features of this novel system (i.e. single-handed pump operation, one-time connection, uninterrupted

SZLYK, SILS, FRANCESCONI, HUBBARD

work, no elevation of canteen, and reduction in number of steps required for drinking) were expected to increase the ease of drinking, and thereby, enhance fluid consumption, particularly during work periods. Responses from the drinking system questionnaire (Table 2) indicate that the FH was received more favorably than the CS. Overall use of the FH was rated as "easy" during both walk and rest cycles while soldiers using the CS rated their system as "slightly hard" when walking and as "neither easy nor hard" when at rest. Moreover, CS subjects rated their system easier to use while at rest than while walking. A 16% higher overall fluid intake rate was noted in FH (0.42 ± 0.11 L/hr), although this was not statistically different from CS (0.36 ± 0.04 L/hr).

Table 2. Drinking System Questionnaire

Group (n)	Overall ease of use during		Ease of water delivery during ¹		Ease of connecting/ disconnecting canteen during	
	Walk	Rest	Walk	Rest	Walk	Rest
<u>Initial response:</u>						
CS (7)	3.6*	2.9*	3.0	2.4**	3.1	1.9
FH (9)	1.9*	1.9*	3.1	1.9	not applicable	
<u>Final response:</u>						
CS (8)	3.4*	3.0*	3.6	2.9**	3.6	2.9**
FH (9)	1.6*	1.5*	2.4	2.1	not applicable	

Values represent average group ratings.-

On a scale from 1 to 5; 1=extremely easy, 2=easy, 3=neutral, 4=hard, 5=extremely hard.

1 Two different questions. For CS: Ease of holding canteen above head and drawing against positive pressure. For FH: Ease of pumping water into mouth.

* Significant difference ($p < 0.05$) between CS and FH.

** Significant difference ($p < 0.05$) between walk and rest.

During the actual exercise periods, subjects using the FH drank 0.35 ± 0.09 L/hr and took an average of 3 drinks per walk cycle while subjects using the CS drank 0.24 ± 0.07 L/hr and took approximately 1 drink per walk cycle. This represents a 46% higher fluid intake rate for FH subjects during the walk cycles, although statistical significance was not achieved due to intersubject variability.

When the fluid intake rates during the rest periods were

SZLYK, SILS, FRANCESCONI, HUBBARD

compared, CS subjects drank, on average, 1.2 ± 0.19 L/hr while FH subjects drank 0.77 ± 0.24 L/hr. These rates of fluid intake during rest were likewise not significantly different between CS and FH because of intersubject variability, although they were 55% higher in CS. However, when comparing the rate of fluid intake while resting to that while walking, both groups demonstrated a cyclic pattern of fluid intake with significantly higher ($p < 0.05$) intake rates during the rest periods than during the walk periods.

Subjects were also asked to rate the taste and temperature of the water because preference (6,13) and perception (13) of beverage temperature may be changed by exercise, hydration status and body temperature. In addition, the first two mouthfuls of water consumed with FH originate from the pumping bulb where it can warm to ambient temperatures between drinks. Since warm water reduces fluid intake (2,9,17), warm water in the pumping bulb may discourage fluid consumption in hot environments. The FH group rated the taste of water as "neither like nor dislike" and water temperature as "neither warm nor cool"; these responses showed no significant change over the course of the trial. These ratings were not different from those rendered by CS, and correlate with the similarity in intake between groups. The rate of water delivery was also assessed and found to be adequate with either system and subjects noted leakage only when the mouth spigot inside the face mask was improperly positioned.

The one-time connection and single-handed operation of the FH proved advantageous to soldiers carrying a rifle especially while walking since they drank greater quantities and with less effort while walking than individuals using the CS. This is an important consideration because periods of inactivity are usually limited during active training or field operations. The higher fluid intake during the walk periods by soldiers using the FH in this study agrees well with observations from a previous study conducted in this laboratory (18). It should also be noted that Kenney (10) reported that to attenuate the effects of hyperthermia and dehydration, it was more beneficial to take smaller, more frequent drinks than to try to compensate for fluid loss in a single drinking episode.

While the one-time canteen connection for the FH group eliminated the requirement for repeated decontamination and attachment of connectors, failure to decontaminate prior to drinking was noted in 22% of subjects using the CS early in the trials (rest periods 1 and 3, respectively). This suggests an increased risk of accidental exposure when using the CS in a contaminated environment. Since the failures occurred early in these trials, they were probably not a result of dehydration or hyperthermia.

Body weight loss was similar in subjects using the CS ($0.34 \pm$

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

0.11 %/hr of initial body weight) and FH (0.37 ± 0.10 %/hr of initial body weight). Sweat rates were also similar at 0.63 ± 0.05 L/hr for CS and 0.67 ± 0.08 L/hr for FH, resulting in equivalent levels of rehydration ($66 \pm 11\%$, CS and $67 \pm 15\%$, FH). These levels of rehydration are greater than those seen in a previous study (18) where only 48% rehydration was achieved by soldiers in MOPP IV configuration (worn over Battle Dress Uniform) and using the CS in a cooler but slightly more humid (29.5°C , 33% relative humidity) environment.

Heat dissipation when working in chemical protective gear may be so severely limited by the minimal evaporation that under even moderate work and environmental conditions, heat storage may be a significant problem for the soldier. Pre-exercise heart rate, rectal temperature (T_{re}) and skin temperature (T_{sk}) were similar in both groups, and rose quickly during exercise. By the final walk period, all three variables had increased significantly over pre-exercise values (Table 3). Subjects using the FH had significantly smaller elevations in T_{sk} and T_{re} than subjects drinking with the CS. Thus, heat storage not only increased significantly from pre-exercise to the final walk in both groups, but the change was significantly greater in CS as opposed to FH subjects (Table 3). These results were not expected because both groups wore the same clothing ensembles, performed the same type of work, and had similar overall fluid intakes. However, the higher thermal strain experienced by the CS group may be due, in part, to the frustration, anxiety and higher metabolic activity associated with this system compared to the FH.

Of the 9 subjects in each group, 4 CS and 2 FH subjects completed the entire 6 hr scenario while the remaining subjects voluntarily withdrew or were removed early from the study because of headaches, dizziness, leg cramps or heart rates that exceeded the safety limits prescribed in the protocol (180 bpm for 5 min while walking). Because no statistical differences in fluid consumption, sweat rates or rehydration were observed between the groups, it is also not surprising that performance was not different. The impact of wearing chemical protective clothing on performance under even moderate work and climatic conditions is demonstrated by the endurance times of 275 ± 28 min for CS and 268 ± 22 min for FH out of a possible 360 min. We have previously reported a 26% reduction in endurance in soldiers working in MOPP IV compared to BDU (18).

Responses to the Environmental Symptoms Questionnaire (ESQ) show that CS and FH were in a similar state of thermal comfort pre-exercise, and experienced increased symptomatology for dehydration and hyperthermia (weariness, bodily aches, headache, dizziness, and thirst) post-exercise (Table 4). However, post-exercise, FH

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

Table 3. Heat strain indices

Variable	CS (n=9)	FH (n=9)
Heart rate (bpm)		
Pre-exercise	88 \pm 3	94 \pm 5
Final walk	161 \pm 5*	161 \pm 6*
Δ HR	73 \pm 6	67 \pm 7
Rectal temperature ($^{\circ}$ C)		
Pre-exercise	37.03 \pm 0.09	37.26 \pm 0.12
Final walk	38.82 \pm 0.14*	38.52 \pm 0.22*
Δ Tre	1.79 \pm 0.14	1.27 \pm 0.15**
Skin temperature ($^{\circ}$ C)		
Pre-exercise	34.23 \pm 0.27	34.55 \pm 0.23
Final walk	36.70 \pm 0.31*	36.46 \pm 0.27*
Δ Tsk	2.47 \pm 0.33	1.91 \pm 0.22**
Heat storage (kcal/m ²)		
Pre-exercise	1224 \pm 18	1224 \pm 28
Final walk	1294 \pm 19*	1270 \pm 30*
Δ S	70 \pm 4	46 \pm 5**

Values are Mean \pm S.E.

Δ = Final walk - Pre-exercise value.

Heat Storage (S) was calculated as (16): $0.84 \cdot (BW) / (BSA) \cdot T_{body}$

Body weight (BW) is in kg

Body surface area (BSA) = $0.00718 \cdot \text{height}^{0.725} \cdot BW^{0.425}$

Body temperature (T_{body}) = $(0.8 \cdot T_{re}) + (0.2 \cdot T_{sk})$

* Significant difference ($p < 0.05$) between pre-exercise and final walk value.

** Significant difference ($p < 0.05$) between CS and FH.

subjects inexplicably noted a higher incidence of symptomatology for bodily aches and light-headedness ($p < 0.01$) than subjects using the CS. This increase in symptomatology is noteworthy since only 6 of the 18 subjects completed the study. In comparison, soldiers dressed in BDU completed the entire 360 min with only minimal increase in heat and dehydration related symptoms (18).

Although subjects from both groups perceived their exertional level during walk periods as progressively more difficult with time, subjects using the FH surprisingly rated their final walk as "very hard" as opposed to those using the CS ("hard", $p < 0.05$). These results may have been due to the higher incidence of bodily aches and light-headedness noted by FH subjects in their ESQ responses.

Under the moderate work and climatic conditions of this study, subjects in MOPP IV configuration (worn over shorts and T-shirt)

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

would have to consume about 0.65 L of fluid per hour to maintain hydration. We previously reported fluid requirements of 0.84 L/hr when chemical protective clothing was worn over BDU (18). These requirements are about 2-fold greater than when wearing BDU (18) and can be reduced by approximately 37% by removing the hooded mask and gloves (MOPP II). These findings emphasize the importance of an efficient through-mask water delivery system.

The higher fluid intake during walk periods, the increased frequency of drinking, and the greater overall ease of use of the FH, indicate the advantage and greater effectiveness of this single-handed, pump operated drinking system, particularly during periods of sustained work.

Table 4. Environmental Symptoms Questionnaire

Complaint	Group	% of Responses Rated as			
		not at all	slight	moderate	severe
		Pre/Post ¹	Pre/Post	Pre/Post	Pre/Post
Weariness	FH	81 / 36	19 / 25	0 / 18	0 / 21
	CS	67 / 47	26 / 24	3 / 6	4 / 23
Bodily aches	FH	89 / 35	8 / 22	3 / 18	0 / 25 *
	CS	86 / 56	7 / 27	2 / 3	5 / 14
Headache	FH	72 / 33	28 / 11	0 / 17	0 / 39
	CS	75 / 56	25 / 17	0 / 6	0 / 22
Light-headedness	FH	91 / 24	9 / 47	0 / 4	0 / 25 *
	CS	93 / 47	7 / 24	0 / 22	0 / 7
Thirst	FH	87 / 38	13 / 29	0 / 18	0 / 15
	CS	90 / 62	10 / 20	0 / 9	0 / 9

Values indicate percentage of responses from all questions about that complaint for which group members rendered that numeric reply.

1 "Pre" reflects responses rendered pre-exercise. "Post" reflects responses rendered upon completing the chamber trial.

* Significant difference ($p < 0.05$) between CS and FH.

UNCLASSIFIED

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

REFERENCES

1. Adolph EF. Physiology of man in the desert. London: Interscience Publ, 1947, pp. 342-351.
2. Armstrong LE, RW Hubbard, PC Szlyk, WT Matthew, and IV Sils. Voluntary dehydration and electrolyte losses during prolonged exercise in the heat. Aviat. Space Environ. Med. 56: 765-770, 1985.
3. Avellini BA. Physiological evaluation of chemical protective clothing. Natick, MA: Navy Clothing and Textile Research Facility Technical Report No. 151, 1983.
4. Borg G. Perceived exertion as an indicator of somatic stress. Scand. J. Rehabil. Med. 2: 92-98, 1970.
5. Boulze D, P Monastruc and M Cabanac. Water intake, pleasure and water temperature in humans. Physiol. Beh. 30: 97-102, 1983.
6. Cabanac M. Physiological role of pleasure. Science 173: 1103-1107, 1971.
7. Du Bois D and EF Du Bois. A formula to estimate the approximate surface area if height and weight are known. Arch. Intern. Med. 17: 863-871, 1916.
8. Durnin JVGA and J Womersley. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women from ages 16 to 72. Br. J. Nutr. 32: 77-97, 1974.
9. Hubbard RW, BL Sandick, WT Matthew, RP Francesconi, JB Sampson, MJ Durkot, O Maller, and DB Engell. Voluntary dehydration and alliesthesia for water. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 57(3): 868-875, 1984.
10. Kenney RA. The effect of the drinking pattern on water economy in hot, humid environments. Br. J. Indust. Med. 11: 38-39, 1954.
11. Pitts GC, RE Johnson and FC Consolazio. Work in the heat as affected by intake of water, salt and glucose. Am. J. Physiol. 142: 253-259, 1967.
12. Rothstein A, EF Adolph and JH Wills. Voluntary dehydration. In: Physiology of Man in the Desert, edited by EF Adolph and

UNCLASSIFIED

SZLYK, SILS, FRANCESCONI, HUBBARD

associates. New York: Interscience, 1947, pp 254-270.

13. Sandick BL, DB Engell and O Maller. Perception of drinking water temperature and effects for humans after exercise. *Physiol. Beh.* 32: 851-854, 1984.

14. Smolander J, V Louhevaara and O Korhonen. Physiological strain in work with gas protective clothing at low ambient temperature. *Am. Ind. Hyg. Assoc. J.* 46: 720-723, 1985.

15. Sohar E, J Kaly and R Adar. The prevention of voluntary dehydration. UNESCO/India Symposium on Environmental Physiology and Psychology, Lucknow, pp129-135, 7-12 December 1962.

16. Stolwijk JAJ and JD Hardy. Partitional calorimetric studies of responses of man to thermal transients. *J. Appl. Physiol.* 21 (3): 967-977, 1966.

17. Szlyk PC, IV Sils, RP Francesconi, RW Hubbard, and LE Armstrong. Effects of water temperature and flavoring on voluntary dehydration in men. *Physiol. Behav.* 45: 639-647, 1989.

18. Szlyk PC, RP Francesconi, IV Sils, R Foutch, and RW Hubbard. Effects of chemical protective clothing and masks, and two drinking water delivery systems on voluntary dehydration. Natick, MA, USARIEM Technical Report No.T14-89, May 1989.

UNCLASSIFIED

George A. Tanton, Dr.,* Charles R. Christensen, Dr., John A. Grisham, Mr.,
and John Stensby, Dr.†
Research Directorate, Research, Development, and Engineering Center, U.S. Army Missile
Command, Redstone Arsenal, AL
35898-5248

Introduction

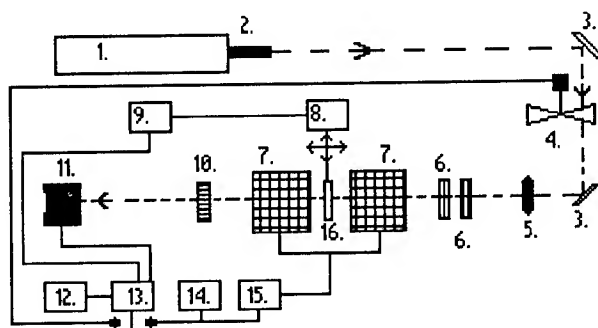
Our objective is to develop innovative noncontact methods that can be applied in ultraviolet (UV) and long wavelength infrared (LWIR) detector manufacturing to increase yield, reduce production costs, and to improve traceability throughout the manufacturing process.

Conventional methods used to characterize semiconductor material do not quickly or efficiently identify areas of a wafer that will produce useful detectors or focal plane arrays. This results in lower production yields and higher unit cost. In the technique described here ^{1, 2} the material of interest is positioned in an amplitude modulated magnetic field and scanned with a linearly polarized laser beam. As the laser beam passes through the material it undergoes Faraday Rotation (FR) proportional to the free carrier concentration in that particular part of the wafer. Since detector device characteristics, such as effective resistance and capacitance, are functions of the free carrier concentration parts of a wafer that will not yield good detectors can be identified in the early stages of manufacturing using this technique. The yield would be significantly increased and the unit cost decreased by identifying and culling wafers or parts of wafers that do not meet specifications at the beginning of the fabrication process. It is expected that this FR technique can be utilized as a universal evaluation tool that will make manufacturing traceability possible by correlating device performance for any part of a wafer with fundamental parameters. This technique is applicable to all major detector materials of Army interest, e.g. mercury cadmium telluride (HgCdTe), cadmium sulfide (CdS), indium antimonide (InSb), platinum silicide (PtSi), and gallium arsenide (GaAs).

Major advantages of this technique are: a) no contacts need be alloyed with the wafer, hence no wafer material contamination. b) surface preparation is not required, c) it is a rapid CAM technique and does not require a skilled operator, and d) automated wafer mapping is

† Permanent Address: Electrical and Computer Engineering
University of Alabama in Huntsville
Huntsville, AL

The bench setup is shown schematically in Fig. 1.



1. LASER, 2 WATTS CW @ 10.6 MICRONS 2. ATTENUATOR 3. MIRROR
4. CHOPPER 5. FOCUSING LENS 6. POLARIZER, ZnSe WIRE GRID 7. SPLIT
ELECTRO-MAGNET, 2.4 KG 8. X/Y TRANSLATION STAGES 9. IBM PC
10. ANALYZER, ZnSe 11. DETECTOR, HgCdTe @ 77 K
12. RECORDER, STRIP CHART 13. LOCK-IN AMPLIFIER 14. SINE WAVE
GENERATOR 15. MAGNET DRIVER POWER AMPLIFIER 16. SAMPLE

Fig. 1. Faraday Rotation, Mapper.

The wafer was probed with a CO₂ laser beam at a wavelength of 10.6 micrometers. The sample was positioned in a magnetic field by X/Y translation stages with a positioning precision of 0.005 mm. The free carrier concentration in a particular part of the wafer was determined by the Faraday rotation as the laser beam passed through the material. The reference plane of polarization was established by two wire grid ZnSe polarizers shown in Fig. 2 and an analyzer, rotated at a bias angle θ was placed after the wafer.

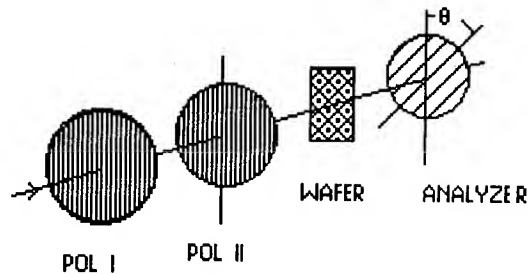


Fig. 2 Faraday Rotation Optical Train

A 25 mm focal length ZnSe lens in front of polarizer I focused the beam to 0.3 mm diameter at the wafer, resulting in a flux density of 10 - 50 W/cm². The split magnet shown in Fig. 3 produced a 10 Hz amplitude modulated magnetic field of ± 2.4 KG. The magnetic field was directed parallel to the propagation axis of the probe beam.

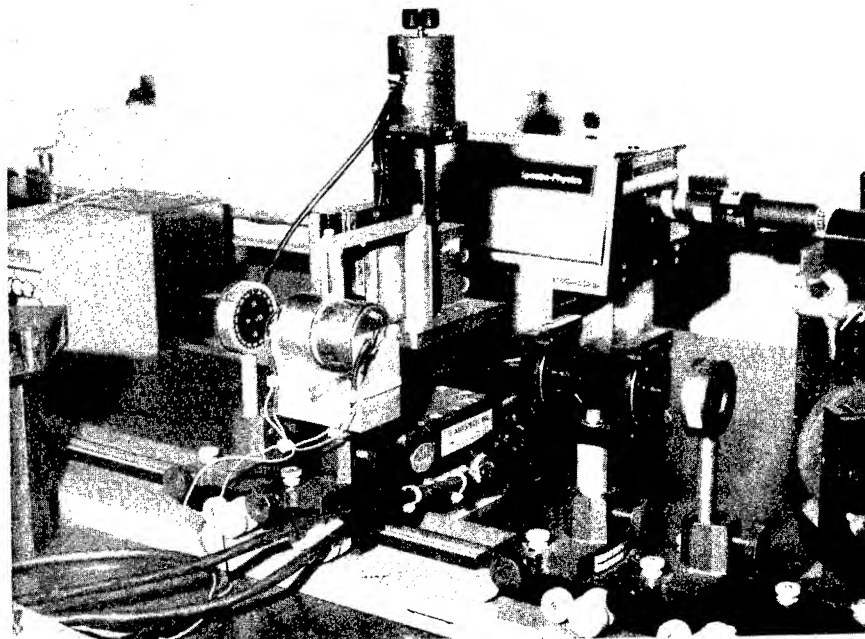


Fig. 3. Split Magnet for FR Measurements

The change in intensity of the probe beam caused by rotation of the plane of polarization (Faraday rotation) and propagation through the zinc selenide Brewster window analyzer was measured by a mercury cadmium telluride detector cooled to 77 K. All system components were located away from any stray magnetic fields that might have affected their performance. A high sensitivity lock-in amplifier, tuned to the modulating frequency of the magnetic field, amplified the detector output and fed it into an IBM PC. The computer was programmed to execute the sequences to control the scanning hardware and collect and reduce the data. Details of the software and theory of operation can be found in another report³.

BACKGROUND THEORY

According to classical theory at wavelengths long compared to the bandgap wavelength. FR due to free carriers is given by.⁴

$$\delta = \lambda^2 \text{Be}^3 \text{NL} / 2\pi c^4 \text{nm}^2 \quad (1)$$

a result that is well established for binary semiconductors. In Eq. (1) δ is the Faraday rotation angle in radians, $e = 4.8 \times 10^{-10}$ esu, the free carrier effective mass m^* was assumed to be 0.015m for MCT and 0.2m for CdS at 300 K, $m = 9.1 \times 10^{-28}$ gm. n = index of refraction of the material, N [cm^{-3}] is the free carrier concentration. $B = 2400$ Gauss, L and λ are the thickness of the sample and wavelength in units of cm, respectively, and π has its usual meaning. The band gap energy, E_g , of CdS is 2.4 eV. It was calculated for MCT from an expression given by Hansen et al.⁵

$$E_g = -0.302 + 1.93x - 0.81x^2 + 0.832x^3 + 0.535 T (1-2x)/1000 \quad (2)$$

where x = percent concentration of Cd. The band gap at $T = 300$ K for long wavelength IR MCT with $x = 0.200$ - 0.220 yields a cut-off wavelength = E_g [eV] / $1.239 < 8$ micrometers. Therefore the output from a CO_2 laser operating at 10.6 micrometers could be used for the probe beam for FR measurements of both MCT and CdS at ambient temperature.

MCT RESULTS

Sample MCT4 was reported to be n-type bulk grown by the travelling heater method and it was investigated in the most detail. The remaining three were supplied as LPE samples. Faraday rotation was proportional to the magnetic field strength B as expected from Eq. 1, Fig. 4 shows the FR signal level plotted against B at an arbitrary position on MCT4. B was determined from the ac voltage across a 1 Ohm precision resistor in series with the magnet coils and from a direct measurement with a Gauss meter.

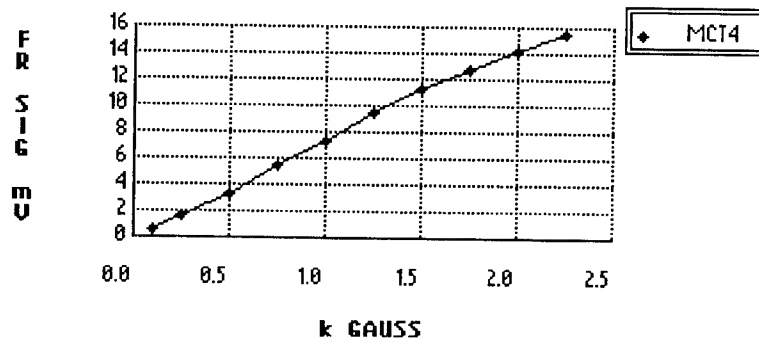


Fig. 4. Faraday Rotation Signal vs Magnetic Field Strength

Fig. 5 shows a comparison of normalized FR as a function of wafer position measured on different days and demonstrates that system repeatability was quite good. Laser fluctuations, although a potential source of scatter in the data, do not appear to be a problem.

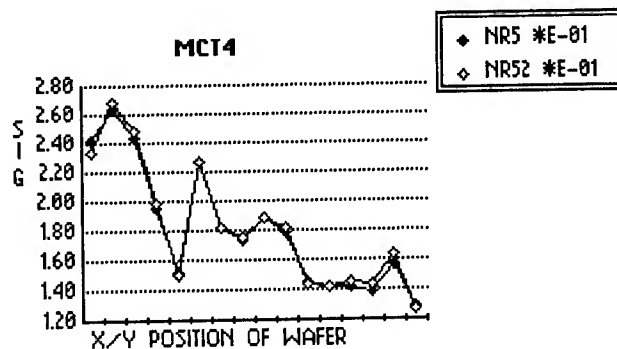


Fig. 5 Faraday Rotation in MCT4 @ 300K. Repeatability Test.

The magnitude of the Faraday rotation angle was measured directly by plotting the detector output as a function of Θ , the angle through which the analyzer was rotated from parallel orientation to polarizer II ($\Theta = 90^\circ$ would be the fully crossed position), Fig. 2. A constant polarization reference plane was maintained for the radiation incident on a sample by rotating the analyzer and keeping the polarizers fixed. Fig. 6 shows the optical train when no sample was in the beam. The beam was chopped at 10 Hz to provide a reference signal for the lock-in amplifier. Fig. 7 shows the resulting detector output as a function of analyzer angle.

The chopper was turned off when a sample was in the modulated field and the amplitude modulated signal generated by Faraday rotation in the sample provided the ac signal necessary for the lock-in amplifier. With sample MCT4 in the modulated field, the detector output had a maximum near 87° , Fig. 8. The difference between the maximum in Fig. 7 corresponds to a FR of approximately 3 degrees.

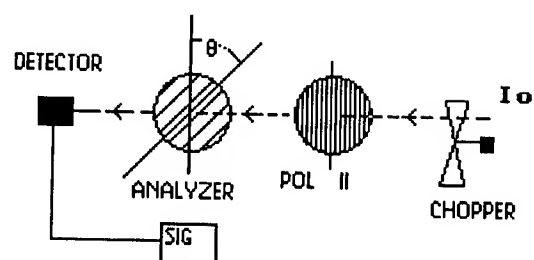


Fig. 6. System Initialization Setup Without Sample

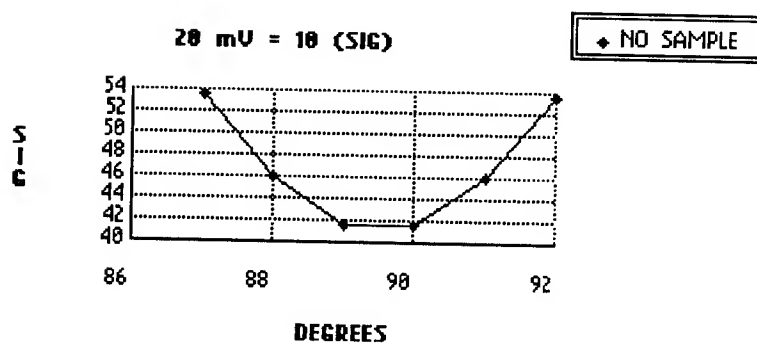


Fig. 7. Detector Output vs Analyzer Angle - No Sample

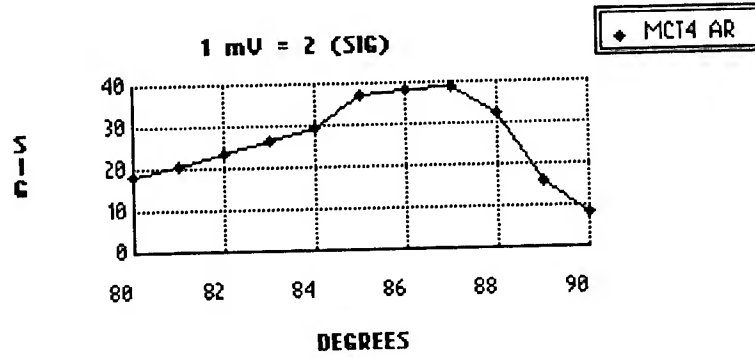


Fig. 8. Faraday Rotation Signal vs Analyzer Angle for MCT4 at 300 K

A more analytical determination of the FR is obtained from the following analysis. From Beer's law and Malus' law the intensity I at the detector is given as

$$I = I_0 \exp(-aL) \cos^2(\theta + \delta) \quad (3)$$

where I_0 = intensity of the laser beam at polarizer II, a = absorption coefficient; L , δ and θ are the wafer thickness, FR angle, and fixed angle between analyzer and polarizer, respectively.

The percent change in I found from Eq. 3 leads to:

$$dI/I = 2\delta (d\delta) \tan(\theta + \delta) \quad (4)$$

A representative value for the magnitude of the FR in MCT4 was calculated from Eq. 1 using the rotation angle determined as follows: A plot of the normalized Faraday rotation vs $\tan(\theta)$, at an arbitrary position on the sample very closely approximates a straight line for θ less than about 80° , as expected from the small δ limit of Eq. 4. The departure from a straight line seen in Fig. 9 for $\theta > 80^\circ$ is due to the contribution of δ .

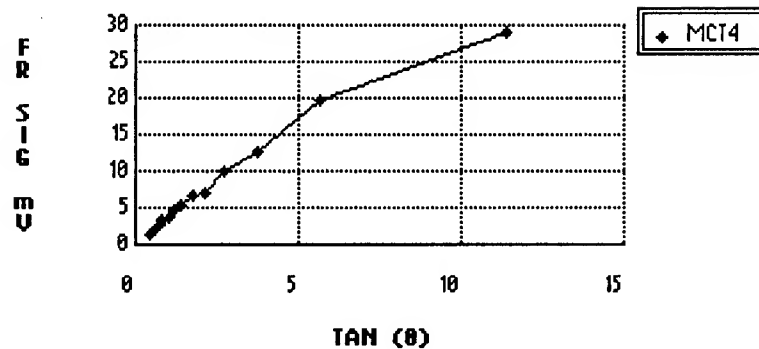
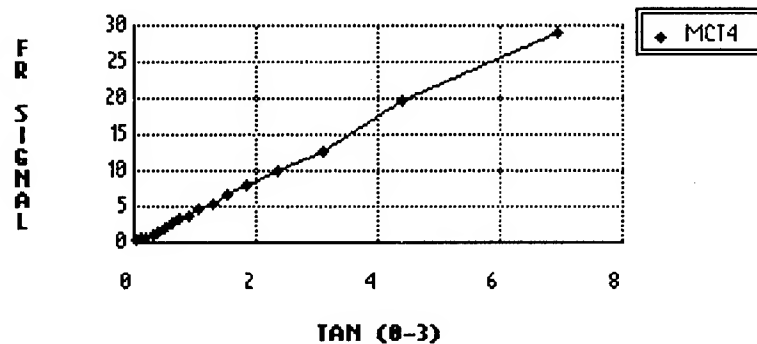


Fig. 9. Faraday Rotation Signal vs Tangent of Analyzer Angle.

The small δ becomes a significant contributing factor as θ approaches 90 degrees because then the system sensitivity to FR theoretically approaches infinity, Eq. 4. In our measurements the gain in sensitivity from setting $\theta = 80^\circ$ instead of 45° was determined experimentally to be equal to 4.5 for both CdS and MCT, compared to the theoretical value calculated from Eq. 4, i.e. $\tan(80)/\tan(45) = 5.7$. The difference between the experimentally determined curve and the extrapolated $\tan(\theta)$ function yields a value of -3 degrees for δ as Fig. 10 shows in good agreement with the value estimated directly from the relatively coarse scale marked in two degree increments on the analyzer holder.

Fig. 10 FR signal vs. Tan ($\theta - 3$).

Solving Eq. 1 for N , here determined at only one arbitrary spot on the wafer, yields $N = 6.7 (10^{15}) \text{ cm}^{-3}$ using $n = 3.55$. This value is somewhat lower than the calculated intrinsic free carrier concentration, $n(i) = 2 \cdot 10^{16}$ for $x = 0.22$ and $T = 300 \text{ K}$, but is consistent with

Hall and FR measurements on CdS at 300 K described later. No attempt was made to measure δ at a position in the wafer corresponding to the calculated mean value of N , only to show that δ obtained as described above is a reasonable value.

The relative free carrier densities were measured in the bulk grown sample MCT4 at room temperature using the automated equipment developed in this laboratory.

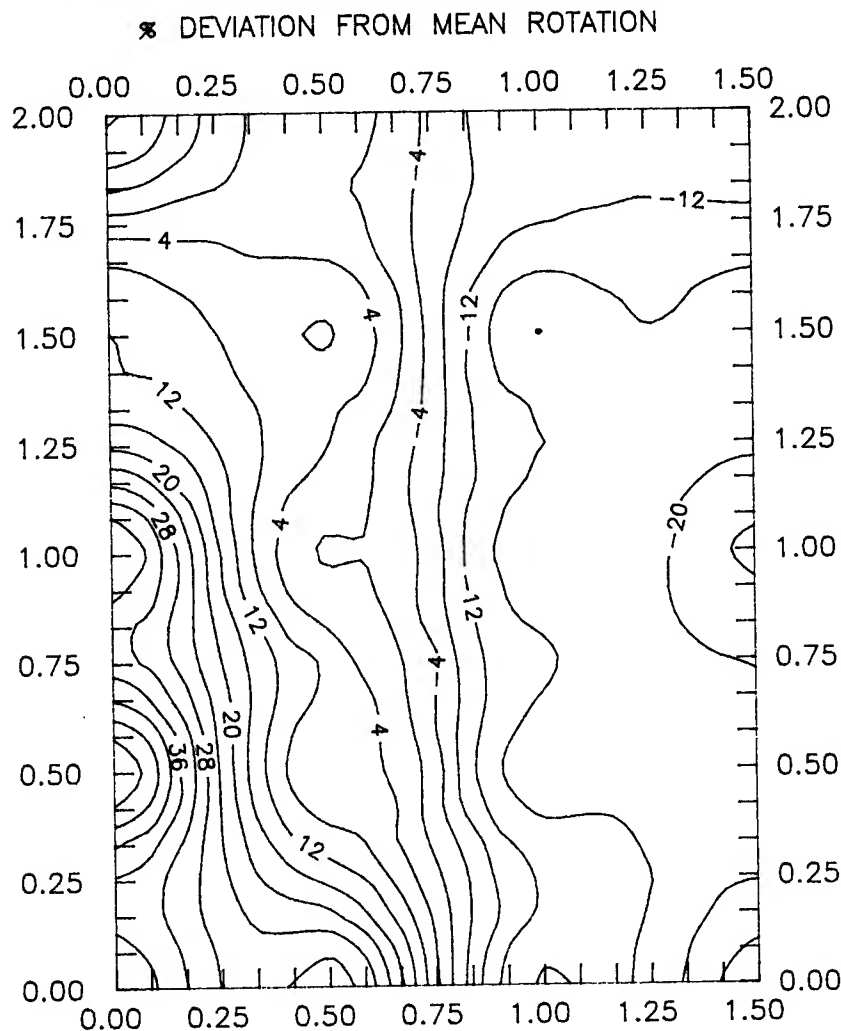


Fig. 11. High Resolution Map of MCT4 (63 Data Points)

Fig. 11 is a high resolution map of MCT4 generated from 63 data points. The acquisition time in both cases was approximately 20 sec/point. The analyzer angle θ was 80° . The free carrier concentration was found to vary 50% or more in sample MCT4. Although it may be desirable to perform measurements on cooled samples Nemiorovsky and Finkman⁷ have pointed out that measurement of the intrinsic carrier concentration at room temperature can be

used to determine the composition and band gap of n-type wafers if the wafer is known to be intrinsic at room temperature.

COMPARISONS WITH LPE RESULTS:

The relative amplitudes of the FR signal in n and p type LPE samples at 300 K are compared in Table I with n-type bulk grown material. Table I also shows the free carrier concentrations at 77 K determined by Hall measurements. Although a direct comparison of FR with Hall results must be made at the same wafer temperature because of possible extrinsic effects.

TABLE I. RELATIVE FARADAY ROTATION (FR) PER CM IN MCT @ 300 K

MCT#	TYPE	THICKNESS	N @ 77 K *10E14/cm ³	FR (mV/cm) ⁺	x
1	p-LPE	18 microns	100	55.5	0. 200-0. 220
2	n-LPE	18 microns	2	30	0. 200-0. 220
3	p-LPE	23 microns	80	40	0. 295
4	n-BULK	864 microns ⁺	2	31.8	0. 200-0.220

Except where marked with +, samples and data in Table I were provided by Michael Grenn, CNVEO, U.S. Army, CECOM

The significant features in this Table are:

1. The Faraday rotation per unit length (FR/cm) increased with increased N as expected.
2. For a given N the magnitude of FR/cm was very close to the same value for n-type LPE and n-type bulk MCT. The FR results were virtually the same for both LPE and bulk grown material even though sample thicknesses differed by a factor of 48. This indicates that the method of growth made no significant difference in the Faraday rotation. Possible extrinsic effects and other factors limit the extent to which one can compare Hall measurements made at 77 K to these FR results made at 300 K. However they appear to be in reasonable agreement.

CdS RESULTS

Faraday Rotation measurements combined with theory and Hall measurements at 300 K give a consistent description of the free carrier concentration over a very wide range of detector and focal plane array material parameters, and demonstrated the applicability of the FR

technique to materials used in the spectral range from UV to IR. This is illustrated by using Eq. 1 to calculate N for MCT4 from Hall and FR measurements on CdS and known values for m^* and n .

Eq. 1 yields the following relationship

$$N_{(MCT)} = \delta_{(MCT)} * [N_{(CdS)} / \delta_{(CdS)}] * 0.009 \quad (5)$$

$$\text{since } 0.009 = (m^{*2}n)_{(MCT)} / (m^{*2}n)_{(CdS)}$$

Room Temperature Hall measurements on two CdS wafers, designated as #7 and #12B, gave free carrier concentrations $N_{(CdS)} = 11.4 * 10^{15} \text{ cm}^{-3}$ and $2.7 * 10^{15} \text{ cm}^{-3}$, respectively. Using Eq. 5 with the maximum and minimum values of FR measured for CdS#7 yields a calculated range of $N_{(MCT)}$ between $2 * 10^{15} \text{ cm}^{-3}$ and $6.5 * 10^{15} \text{ cm}^{-3}$. The maximum and minimum values of FR in CdS#12B yield a calculated range of $0.9 * 10^{15} \text{ cm}^{-3}$ to $6.7 * 10^{15} \text{ cm}^{-3}$ for $N_{(MCT)}$. These values are consistent with the value of $6.7 * 10^{15} \text{ cm}^{-3}$ calculated for MCT4 using $\delta_{(MCT)} = 3^\circ$ measured at a single point as previously described.

CONCLUSIONS

Major conclusions from this effort are summarized as follows:

- a. The first high resolution mapping using the FR technique to determine free carrier concentration uniformity of MCT as a function of position of a wafer was demonstrated.
- b. Qualitative results of FR are consistent with those reported for LPE and bulk grown n-type material.
- c. Quantitative consistency between results of Hall measurements, FR, and theory was demonstrated over a wide range of material parameters covering the spectral range from UV to IR.
- d. In the limited circumstances of this first MCT FR mapping demonstration, the FR method appears to be a viable technique applicable to semiconductors that operate anywhere between the UV and LWIR. FR measurements at cryogenic temperatures on additional samples are needed to demonstrate the full capability of this technique.

ACKNOWLEDGEMENTS

The cooperation extended by the Center for Night Vision and Electro-Optics (CNVEO), CECOM, by providing the MCT samples and Hall data used in Table I, taken by Michael Grenn (CNVEO), is acknowledged. The work reported here was made possible through the support of the Systems Engineering and Production Directorate and Research Directorate, MICOM. Technical assistance of Rachel Wright is gratefully acknowledged.

REFERENCES

1. M. Balkanski and E. Amzallag, *phys. stat. sol.* 30, 407 (1968).
2. G. A. Tanton, J. A. Grisham, C. R. Christensen, and S. Razi, "Proc. Electronics and E-0 Materials for Smart Munitions Workshop", p. 105-114, 13-14 May 1987, Redstone Arsenal, AL, GACIAC PR-87-01.
3. John Stensby, C. R. Christensen, G. A. Tanton, and J. A. Grisham, "A Computer Controlled System for Measurement of Faraday Rotation in Semiconductors," U. S. Army Missile Command Technical Report TR-RD-RE-90-5, 1990.
4. M. J. Stephan and A. B. Lidard, *J. Phys. Chem. Solids*, 9, 43 (1958).
5. G. L. Hansen, J. L. Schmit, T. N. Casselman, *J. Appl. Phys. (USA)* vol. 53 #10 (1982) p. 7099-101.
6. John Brice and Peter Capper, *Properties of Mercury Cadmium Telluride*, INSPEC, The Institution of Electrical Engineers, The Gresham Press, 1987.
7. Y. Nemirovsky and E. Finkman, *J. Appl. Phys. (USA)* vol. 50 #12 Dec 79, p. 8107-11.

A Determination of the Intermolecular Potential of Nitromethane

* S. F. Trevino, Dr.
U.S.Army Research Development and Engineering Center
Picatinny Arsenal NJ 97896

and
National Institute of Standards and Technology
Gaithersburg MD 20899

and

Betsy M. Rice, Dr.
U.S.Army Ballistics Research Laboratory
Aberdeen Proving Ground MD 21005

Introduction

The molecule nitromethane $\text{H}_3\text{C}-\text{NO}_2$ is, by standards of organic chemistry, a very simple system. One property which might rank it as particularly interesting to a physicist is the ease with which the end groups of the molecule, the methyl group CH_3 and the nitro group NO_2 , are able to rotate about the C-N bond with respect to each other. The atoms which comprise these two groups are bound to each other with very strong chemical bonds. These bonds result from the sharing or exchange of a significant number of electrons. This intimate sharing of electrons among the nuclei comprising the chemical groups results in forces which tend to retain the overall geometric configuration of these groups. In the molecule, these two groups are bound to each other through the chemical C-N bond which is similarly a result of electron sharing between the carbon and nitrogen atoms. The electrons producing this particular bond arrange themselves in a very symmetric configuration about the line joining the C and N atoms. This configuration of itself will not provide a force restricting the rotations of the end groups about this bond. The only other source of such a force is that experienced by the interaction between the atoms (the hydrogens) of one end group with the atoms (the oxygens) of the other. These atoms do not share electrons between each other so as to produce chemical type bonds with the result that the interaction forces are substantially weaker. In addition because of the arrangement of these atoms with respect to each other, the net force on the "rigid" methyl group preventing rotation about the C-N bond very nearly cancel. Indeed this has been verified by measurements of the rotational motion with microwave

spectroscopy¹. It is this fact which makes this system so potentially interesting in that in the solid state the properties of the rotational motion of the methyl group will be determined entirely by the interactions produced by neighboring molecules. The rotation of the nitro group is substantially diminished because of its significantly larger moment of inertia.

In the following sections are described the experimental measurements by which the crystal structure and methyl group rotational properties of nitromethane in the solid have been determined. In the last section is described the theoretical methods used which resulted in the determination of the intermolecular potential which is consistent with these measurements.

The Crystal Structure¹

Nitromethane melts at ≈ -40 °C. It was probably for this reason that no attempt had previously been made to determine the crystal structure of solid nitromethane. A knowledge of the structure is *essential* to ensure a correct interpretation of any measured quantity which is to be related to some phenomena on an atomic range. The structure is determined by a study of the scattering of an appropriate radiation from the material. The most prevalent radiation used is X-ray which is scattered by the electrons surrounding the nuclei. This latter fact makes this method most sensitive to the location of those atoms with the most electrons. A second radiation in common use is the neutron (when a source of neutrons is available). The neutron is scattered by the nuclei and is in general equally sensitive to all nuclei irrespective of the number of electrons associated with them (there are exceptions to this rule). Thus an accurate determination of the positions of the light atoms (hydrogen) in the presence of heavier atoms would be greatly aided by the use of neutron scattering. In the present case a combination of both techniques was used in the successful determination of the crystal structure. X-Ray diffraction from a single crystal grown *insitu* on the diffractometer and held there at -45 °C was used to determine the space group ($P2_12_12_1$) and the positions of the heavier atoms of nitromethane. Neutron diffraction from a powder of (for technical reasons) the deuterated molecule was used to determine the positions of the deuterium atoms (which replaced the hydrogen atoms). These were accomplished at temperatures of from 4 K to 150 K. It was required to use these two complementary methods because the crystal size necessary for the X-Ray study is more than an order of magnitude smaller, and therefore easier to produce, than that required for neutron diffraction. Given the results of the X-Ray study, the results of a neutron scattering from a polycrystalline sample were more than adequate to locate the deuterium atoms. The attainment of very low temperatures is also more easily accomplished with a powder sample on the neutron diffractometer. In Fig. 1

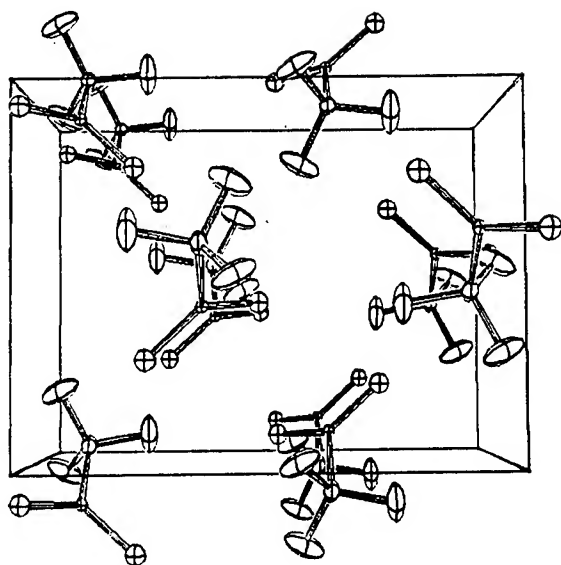


Fig. 1 The crystal structure of nitromethane-d₃

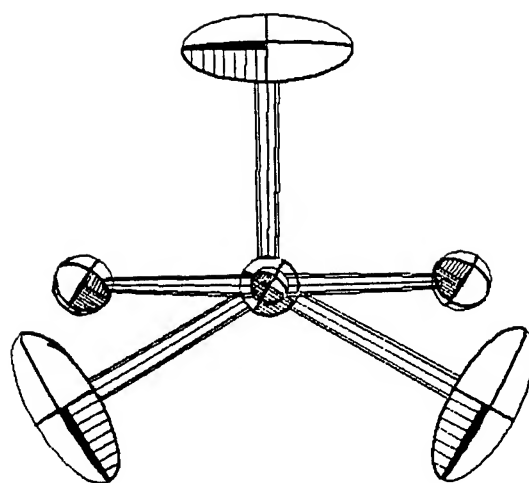


Fig. 2 The molecule of nitromethane-d₃ as it occurs in the crystal as viewed along the C-N bond. The ellipsoids represent the thermal motions of the various atoms.

is presented the molecules of nitromethane as determined by the neutron diffraction at 4 K. Note that all the atoms are well located and that the figure implies that the location of all neighbors of the methyl hydrogens are available. Fig. 2 presents a view of one molecule with the C-N bond perpendicular to the plane of the paper. The three atoms represented by large quasi ellipsoids related by three fold symmetry are the deuterium/hydrogen atoms and the two by smaller circles at the ends of an almost straight line are the oxygen atoms at the other end of the molecule. A rotation of the methyl group about the C-N bond occurs between equilibrium positions separated by 120° . This will be important in subsequent discussions of experimental measurements concerning this motion. In addition, the large extent of the ellipsoids, which represents the distribution of the deuterium atoms at this temperature (4 K), has further implications on the properties of the rotational motions as will be discussed below. These measurements also indicate that there is no drastic change in the crystal structure between the lowest temperature, 4 K, and the highest ≈ 230 K measured (except for changes in the lattice constants). Finally, a most important result of this study is that all the molecules in the crystal are related to each other by the symmetry operations required by the space group. The physical consequence of this statement is that all the methyl groups find themselves in exactly the same environment of neighboring molecules so that what ever description applies to one applies to all. We are thus, fortunately, searching for one only description of

the rotation of the methyl group.

Quasi-Elastic Neutron Scattering²

If the kinetic energy of motion, and in particular of rotational motion, of the methyl group is large enough to induce the methyl group to overcome the barrier between the equilibrium positions, then the methyl group may occasionally re-orient about the C-N bond so that the hydrogen atoms interchange positions by an angle of 120°. If such motion occurs in the solid, it can be detected by its effect on a neutron of a given energy as the neutron traverses the sample and experiences a scattering by the hydrogen nuclei. A simple model which can be used to describe this effect is one in which the methyl group remains in its equilibrium position describing relatively small excursions for a time τ after which it performs an instantaneous jump between adjacent equilibrium positions separated by 120°. In this model it does not matter whether the rotation is in one direction or the other (the effect of the direction being undetectable) but only that the time between jumps be long compared with the time for the jump. Such a model predicts that the energy distribution of the scattered neutrons is described by the following equation.

$$I(E) \propto [1+2\sin(Qd)/Qd] \times \delta(E-E_0) + 2[1-\sin(Qd)/Qd] \times \frac{3/2\tau}{(3/2\tau)^2 + (E-E_0)^2} \quad (1)$$

Here E is the energy of the scattered neutron, E_0 is the energy of the incident neutron, Q , the momentum transfer, is a known function of E , E_0 and the scattering angle, d is the distance between equilibrium positions and τ is the residence time described above. There are two terms in this equation. The first describes a scattering in which the neutron does not change energy but only direction and the second describes the energy spectrum of the scattered neutrons as a Lorentzian whose width is determined by the residence time τ and whose maximum is centered at the energy of the incident neutron (hence the name quasi-elastic). Note that the relative intensity of the two components is prescribed by the geometry of the process i.e. the jump distance d and the scattering angle through the function Q . If the model is capable of an adequate description of the scattering process, one has in hand a method of measuring τ which is presumably determined by the barrier presented to the rotation. Neutron spectrometers capable of such a measurement exist on the research reactor of the National Institute of Standards and Technology (NIST) and are here used for this purpose. Fig. 3 presents the results of such a measurement for a powder of CH_3NO_2 at a temperature of 78 K. The solid line is a quite adequate fit to the data using the model described above in which the resolution of the instrument is folded into the model predictions. The

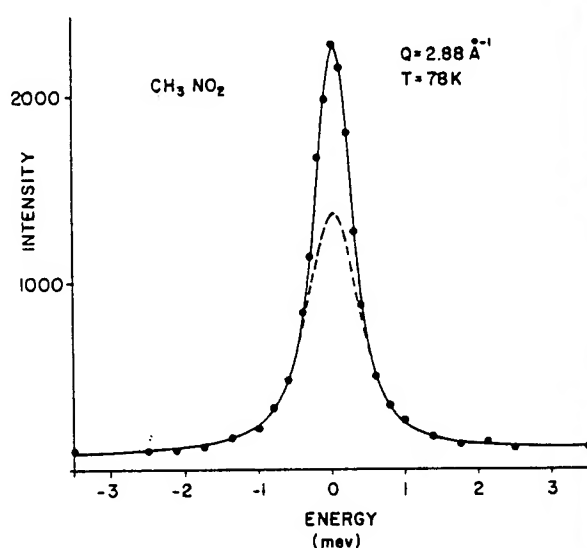


Fig. 3 Quasi-elastic neutron scattering spectra of nitromethane at 4 K. The solid dots represent the data, the solid line the fit to the data and the dashed line the contribution to the spectrum due to the Lorentzian with the model discussed in the text.

dashed line describes the contribution due to the Lorentzian term. From this data a residence time of 3.1 picosec is obtained at this temperature. The residence time should be temperature dependent for it is the temperature which determines the kinetic energy available for rotational motion which energy allows the barrier to be overcome. It can be shown that the temperature dependence is given by

$$\tau(T) = \tau_0 \times \exp(E_b/kT) \quad (2)$$

Where τ_0 is the residence time at infinite temperature (i.e. the minimum residence time) and E_b is the activation energy which must be overcome in order to produce a reorientational jump. Measurements at several scattering angles have been performed in order to adequately test the model (it proves

quite adequate) and at temperatures of from 50 K to 150 K to obtain E_b . A value of E_b of 230 cal/mole (10 meV or 116 K) is obtained. This constitutes one measure of a property of the environment which determines the rotational motion of the methyl group.

Inelastic Neutron Scattering³

At very low temperatures (4 K) the methyl group has insufficient energy to overcome the barrier and instead remains in its equilibrium position (thus using the parameters determined for Eq. 2, the residence time at 4 K is ≈ 1 sec). At these positions however it can still execute small rotational like vibrations about the equilibrium positions. Because the mass of the rotating particles and the distances through which they move are so small, quantum mechanical considerations operate. In particular, these considerations require that the group may sustain rotations whose energies are characterized by discrete (quantized) values. The values of these energies are completely determined by the moment of inertia of the rotor and the shape of the potential well in which the rotors reside. There is a minimum energy level below which the rotating

group may not exist, referred to as the ground state. Above this level there exist at least one if not more levels whose energies lie below the top of the potential well, these are referred to as bound excited levels. The number of bound levels is a function of the height of the potential, there being more as the height of the potential increases above the ground state. An additional quantum effect, tunneling, affects each of these levels. Even though there is not enough thermal energy for the methyl group to overcome the barrier and perform reorientations by jumping over it, the effect of quantum tunneling allows the methyl group to reorient by passing through the barrier. This is purely a quantum effect and is allowed by the "uncertainty principle". The effect of this phenomena is to split each of the energy levels discussed above into two closely spaced levels. The amount of splitting is an extremely sensitive function of the shape of the potential at the position of the levels. As mentioned above, the configuration of the energy level scheme is also a function of the moment of inertia. Additional information could be obtained if one could change this quantity without changing the potential. This is readily accomplished by replacing the hydrogen atoms with deuterium atoms.

Now, neutrons are scattered by the nuclei which make up the material and if the nuclei are in motion during the scattering, energy may be exchanged between the neutron and the moving nuclei. Again, quantum mechanics requires that the energy exchange during the scattering, when it occurs, be such that the change in the energy of the rotating body be exactly the difference between two of the allowed energy levels, the difference exhibiting itself as an energy change in the scattered neutron. Neutrons are not the only radiation which can accomplish such transitions (for example, a laser beam also causes these energy transfers) but in the present case of rotational motion, the neutron is by far the most sensitive probe available. Neutron spectrometers of appropriate energy resolution and sensitivity located at the NIST research reactor in Gaithersburg MD and at the Institute Laue-Langevin at Grenoble France have been used in these measurements. In Fig. 4 is presented inelastic neutron scattering data obtained for both CH_3NO_2 and CD_3NO_2 at 4 K with two spectrometers, one sensitive to the energy range in which the tunnel splitting of the ground state occurs, and the other to the energy range at which the excited states exist above the ground state. In three of the spectra the assignment of the observed features to the rotational energy levels is clear. Only in the spectrum for CH_3NO_2 in the energy region near 7 meV where there appears

to be two unresolved peaks is the assignment not unambiguous. The question at this point is whether both of these features are associated with the rotational excited states or not. In order to clarify this question, the measurements were repeated with the sample under hydrostatic pressure. It was chosen to perform this measurement in the hope that features due to the rotational state would be more sensitive to changes caused moderate pressure than would be structure due to other (translational type e.g.) motions. Fig. 5 is presents the data obtained at a pressure of .5 GPa

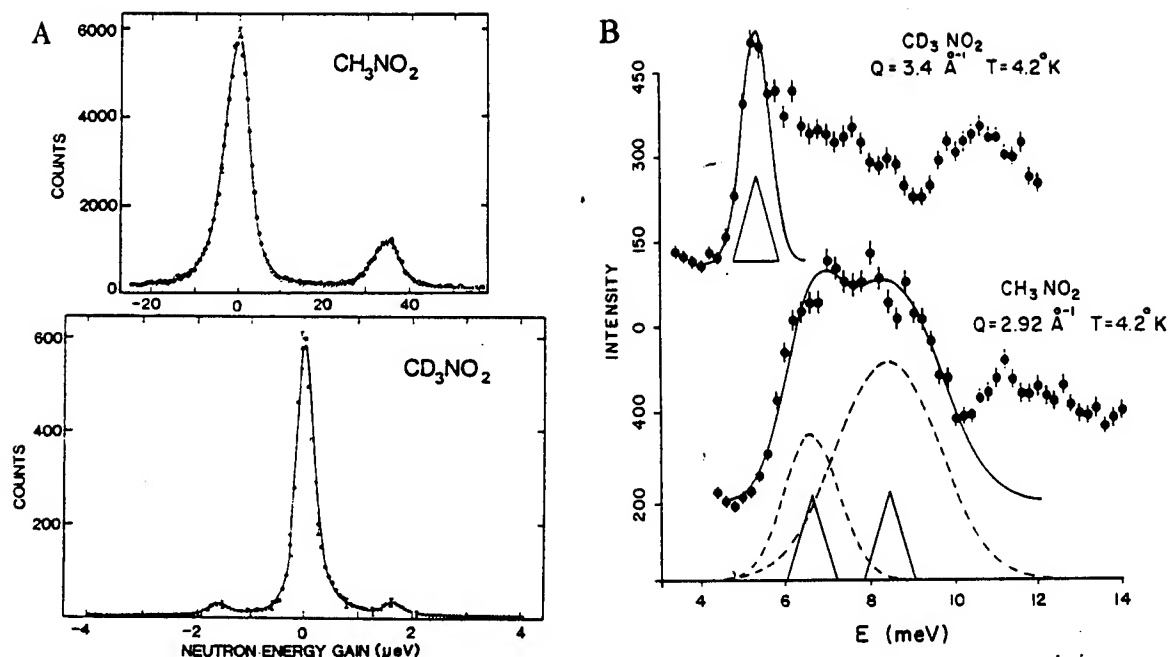


Fig. 4 Inelastic neutron scattering data of CH_3NO_2 and CD_3NO_2 at 4 K in a) the energy region of the tunnel splitting of the ground state and b) the energy region of the excited states.

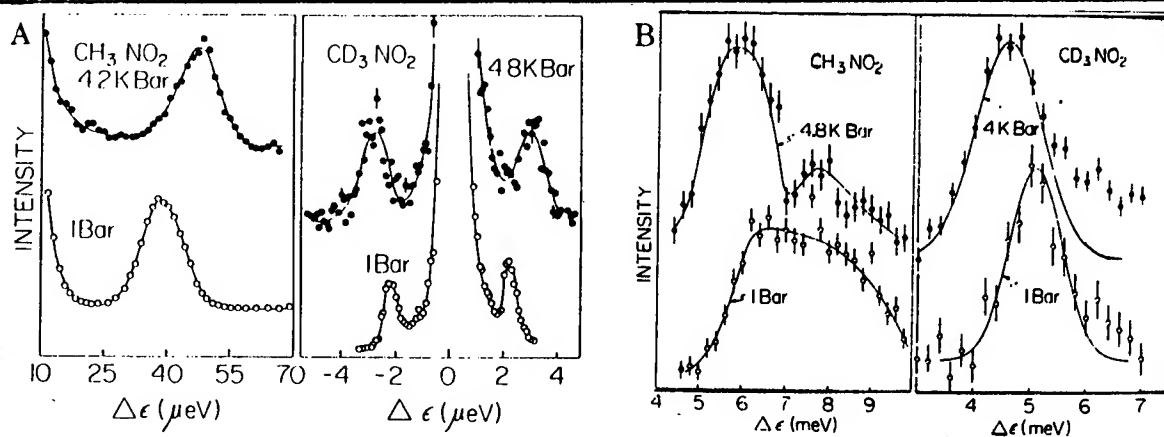


Fig. 5 Inelastic neutron spectra of CH_3NO_2 and CD_3NO_2 at 4 K as a function of pressure in the two energy regions of a) ground state tunnel splitting and b) excited state.

(5000 atmospheres) for both samples in the same two energy regions and compared with the original data of Fig. 4. Note that in the case of CH_3NO_2 in the region of energy of the first excited state (5-7 meV) it is now clear that only the low energy feature indeed moved with applied pressure.

Note also that the corresponding peaks in the two isotopic samples move in the same direction with pressure confirming the assignment of these peaks as due to the same source, viz the rotational motion. The tunnel splitting peak in both moves to higher energies and the energy of the first excited state changes to lower energy upon the application of pressure. This observation was a complete surprise. At the time, there existed eight other systems in which methyl rotations had been investigated as a function of pressure. In all these systems, the change of all the energies with pressure had been measured to be in exactly the opposite direction. It was quite clear that the interpretation of the previous studies, viz that the potential barrier simply increased with pressure, could not be made here. But then neither could one propose that the barrier would decrease with pressure (this surely be bizarre). The impasse was resolved when it was realized that a change in the shape of the potential could produce such effects. The fact that the equilibrium positions of the hydrogens are separated by 120° implies that the potential must have at least 3-fold symmetry with respect to rotation about the C-N bond. Thus an expansion with respect to the rotation angle can be expressed as:

$$V(\theta) = V_3/2 \times [1 - \cos(3\theta)] + V_6/2 \times [1 - \cos(6\theta + \delta)] + \dots \quad (3)$$

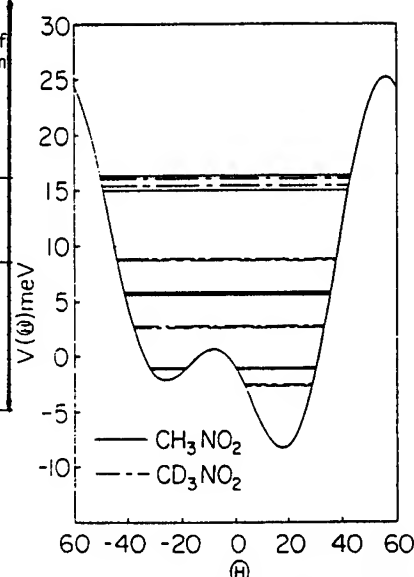
where θ is the rotation angle about the C-N bond and V_3 , V_6 and δ are parameters. In the above mentioned eight "normal" studies the value of V_6 was zero and the potential had a very simple shape. In the present case it was found that the inclusion of a substantial amount of V_6 was sufficient to explain all the data at a given pressure and that the effect of pressure could be accounted for by an increase of V_6 over V_3 . In Table 1 is given the calculation of all the observed energy levels using this potential and compared with the observations. This table presents measurements made from spectra in addition to those specifically presented here. In Fig. 6 is plotted the potential resulting from the parameters appropriate to the ambient pressure data and the calculated energy levels for both isotopic species. Several points are worthy of special comment. A given set of values for the parameters V_3 , V_6 and δ are sufficient to reproduce adequately the observations of the energy levels for both isotopic species. The potential indeed does not change upon isotope substitution. The value of δ is not zero, this implies that the source of this term is not located symmetrically about the equilibrium positions of the methyl group. A substantial region of the potential is sampled by the measurements increasing the confidence that the shape is well described by this function. The pressure dependence is well accounted for by an increase in the relative magnitude of V_6 viz a viz V_3 . A microscopic explanation should account for all these features.

Table 1.

Measured and calculated energy levels of the two isotopic species of nitromethane at two pressures and 4 K. The parameters in the calculation are for Eq. (3) and are: 1 bar, $V_3=25.5$ meV, $V_6=-15.5$ meV and $\delta=30^\circ$; 5 kbar, $V_3=22.5$ meV, $V_6=-17.2$ meV and $\delta=25^\circ$.

	Ground State splitting (μ eV)		First excited state (meV)		Second excited state (meV)		Pressure
	Meas.	Calc.	Meas.	Calc.	Meas.	Calc.	
CH_3NO_2	35	34	6.7	6.9	17.5	16.7	1 bar
CD_3NO_2	1.7	1.3	5.3	5.2	10.6	11.4	1 bar
CH_3NO_2	54	54	5.6	5.6	15.8	5 kbar
CD_3NO_2	2.8	2.5	4.4	4.1	10.7	5 kbar

Fig 6. The rotational potential of the methyl group in nitromethane corresponding to the data at 4 K and ambient pressure and the energy levels for the two isotopic species as noted in the figure.



As mentioned above, the crystal structure determination allows the location of the atoms neighboring the hydrogen atoms of the methyl group. It is presumably interactions between these atoms and the hydrogen atoms which give rise to the rotational potential. A calculation of $V(\theta)$ can be accomplished in the "mean field approximation" in the following manner. Without changing the position of the neighbors, at each value of θ , calculate the total potential energy resulting from the interaction of each of the three hydrogens atoms with each of their neighbors. One only needs the potential of interaction of the hydrogens with their neighbors as a function of the distance to these neighbors. Several explicit functions for such a description have been suggested in the past based on extrapolations from basic physical principles. One of the most common in use is the Lennard-Jones potential. It is of the form:

$$v(r) = \epsilon [(\sigma/r)^{12} - (\sigma/r)^6] \quad (4)$$

where r is the distance from the hydrogen to the neighbor in question and ϵ and σ are parameters. The first term describes the repulsive and the second the attractive nature of the interaction. The sum leads to a minimum which presumably has physical significance. For several physical reasons, we believe that the hydrogen oxygen interaction is the most important in the present case. Indeed we find that for $\epsilon = .283$ eV and $\sigma = 2.18$ Å, the $V(\theta)$ given in Fig. 6 is exactly reproduced. With these values for $v(r)$, it becomes clear that the source of the V_6 term in $V(\theta)$, that is the "bump" at the bottom of the potential well, is due to the interaction of the nearest H-O neighbors which interaction is due to the repulsive part of $v(r)$. The neighbors at larger distances experience an attractive interaction. This intermolecular interaction is therefore also consistent

with the pressure dependence of $V(\theta)$ in that as one applies pressure, the nearest neighbors experience an even larger repulsive interaction, for they are forced closer to the hydrogens, thus causing an increase in V_6 viz a viz V_3 . Finally, the asymmetric location of the "bump" (the non-zero value of δ) is a consequence of the spatial distribution of the oxygen atoms surrounding the methyl hydrogens. All of the requirements for a successful microscopic explanation are in hand.

Recall from the second section that the activation energy was measured with quasi-elastic neutron scattering to be ≈ 10 meV. Inspection of the potential of Fig. 6 however reveals that the ground state is almost 20 meV below the top of the potential well. It is clear that although the activation energy and the potential barrier are related, in this case there is not a one to one correspondence between their values. It is to be noted here that there are many systems for which both of the above quantities are obtained. The two values are indeed identical for the great majority of cases in particular for those for which the activation energy is substantially larger (by at least a factor of 2) than that encountered here. For those systems in which the activation energy is comparable to or smaller than that found here, it is systematically smaller than the barrier required for the energy levels. This nonlinear relationship for small barriers seems to be universal. It is interesting to speculate on the origin of the difference so obvious in the case of nitromethane. Remember that the $V(\theta)$ obtained is due to spectroscopic data for which the sample was maintained at 4 K whereas the activation energy was obtained from data for temperatures between 50 and 150 K. It might be that thermal motion caused by the elevated temperatures results in a change in $V(\theta)$. Alternately if this is not the case, some interaction with other motions which the molecules can execute, such as translations or rotations of the entire unit, might facilitate tunneling when the methyl group finds itself in one of the rotational excited states. One method which suggests itself to shed light this question is that of molecular dynamics. For this method, one simulates the motions of the molecules with time at a given temperature and pressure by solving the classical equations of motion as a function of time. One needs only a reliable interaction potential (which we thought we had) and a fast computer (and a good programmer). In preparing for this project, we noticed that the minimum of $V(\theta)$ occurred some 20° away from 0° . This caused some concern in that the value of 0° for θ is not arbitrary, it is the value of the equilibrium orientation of the methyl group as determined by the crystal structure studies. We are in some difficulty.

Determination of the Intermolecular Potential

The form for the interaction potential $v(r)$ given in Eq.4 is not the only one in use. A search for a more appropriate function was undertaken, the requirement being that the shape of $V(\theta)$ be obtained (the shape is

determined by the measured energy levels) but that the minimum occur at a value of $\theta = 0^\circ$. Also the mean field approximation was relaxed, i.e. the neighbors were allowed to move to a certain extent during the calculation of $V(\theta)$. All to no avail. The minimum in $V(\theta)$ could not be budged away from $\approx 20^\circ$. It proved useful to write explicitly the expression used in the calculation of $V(\theta)$. It is:

$$V(\theta) = \sum_j v[r_j(\theta)] \quad (5)$$

where the sum is over the j neighbors of the hydrogens and $r_j(\theta)$ is the distance of these j neighbors from the hydrogens at each value of θ . The object of the search was for the form of v to produce the correct properties of V . The forms used to this point proved inadequate. It was clear that the requirements of geometry imposed by the crystal structure must be allowed to lead to the desired result. For this purpose Eq. 5 is rewritten as:

$$V(\theta_i) = \sum_k \Phi_{ik} v_k \quad (6)$$

where the Φ are defined as:

$$\Phi_{ik} = N \sum_j \delta[r_k - r_j(\theta_i)] \quad (7)$$

This is essentially a (properly normalized by N) pair distribution function at a given value of θ_i which when multiplied by the corresponding value of v_k will produce a value of $V(\theta_i)$. The only approximation used in deriving Eq. (6) from Eq. (5) is that the value of v_k is constant for the range of r from r_k to r_{k+1} . The number of values of i and k are somewhat arbitrary. A study was conducted to determine a "reasonable" number of i which would define the shape of $V(\theta)$ adequately, and the number of k which would result in a reasonable relationship between v and V . Written in this form, however, the information concerning the geometry is contained in the function Φ explicitly and the v_k are parameters to be determined without recourse to a specific functional form. This is not a problem amenable to a least square solution.

Problems similar to this are encountered in for example astronomy and image enhancement. A certain number of observations are in hand from which is to be extracted information concerning a larger number of parameters describing a model some how related to reality. In the case of image enhancement, for example, the image which produced a photograph of inadequate quality, usually due to poor resolution, is sought. One only has the photograph. A method is sought which, in an unbiased manner, can

change the values of the parameters (the light and dark shades of the image sought) from an initial distribution to one which is the most likely consistent with the observations.

It has been shown⁴ that the only consistent selection procedure for the set of parameters, v_k in our case, is one which maximizes the entropy function

$$S = -\sum_k v_k \ln(v_k/v_k^*) \quad (8)$$

subject to the constraint that the parameters be consistent with the set of values of $V(\theta_i)$. This constraint is expressed as the requirement that the statistic

$$\chi^2 = 1/N \sum_i [V(\theta_i) - V^*(\theta_i)]^2 \quad (9)$$

have the expected value of 1. Here $V^*(\theta_i)$ are the values obtained from the set of parameters v_k^* , N is the number of values of i and $V(\theta_i)$ represents the rotational potential obtained above. Obtaining the conditions stated above results in the prescription that a new set of parameters v_k is to be obtained from the previous set v_k^* by the formula

$$v_k = v_k^* \text{EXP}\{\sum_i \phi_{ik} \times [V(\theta_i) - V^*(\theta_i)]\} \quad (10)$$

In an iterative manner, we begin with some initial set of v_k^* and obtain new values until some convergence criteria is met. In our case we begin with the set of v_k corresponding to the values for the Lennard-Jones potential given above. The result of this exercise is shown in Fig 7. Here is plotted the original $v(r)$, that produced by the entropy maximization and the difference. Recall that the entropy maximization involved the approximation that v_k represented some average and constant value of $v(r)$ between r_k and r_{k+1} . The results of this procedure were thus interpreted as a guide for changes in the original function $v(r)$ which would produce the desired results. Inspection of Fig. 7 indeed provides a powerful guide for the use of additive terms to the Lennard-Jones potential which would result in the desired results. Several Gaussian functions were used to, in a least square fit, obtain finally the potential presented in Fig. 8 which figure contains the original $v(r)$, the final $v(r)$, the suggested changes from Fig. 7 and the final changes (the details of the exact parameters for the functions which resulted in this potential are

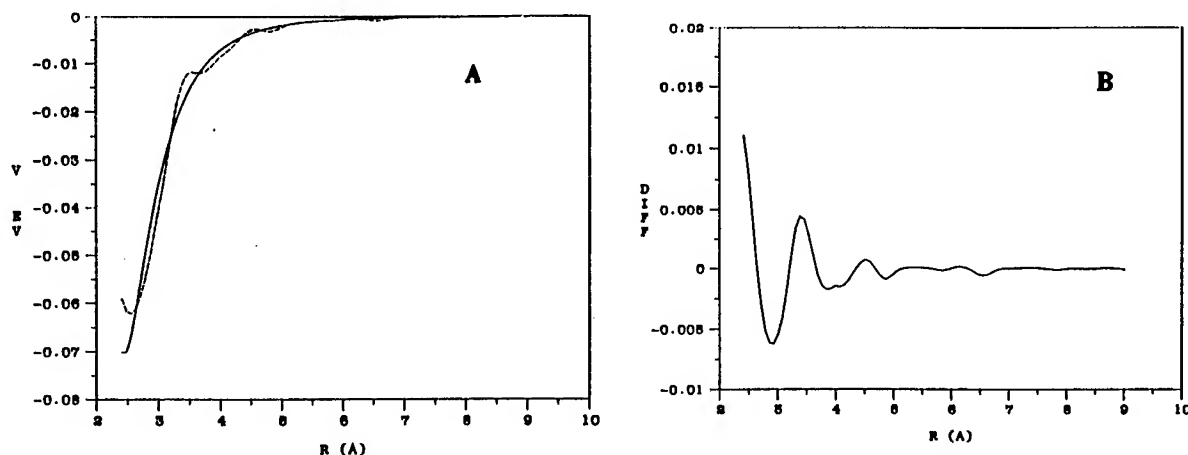


Fig. 7 The potential for the H-O interaction as a function of separation for a) the original Lennard-Jones potential (solid line), that produced by the maximum entropy exercise (dashed line) and b) the difference, as discussed in the text.

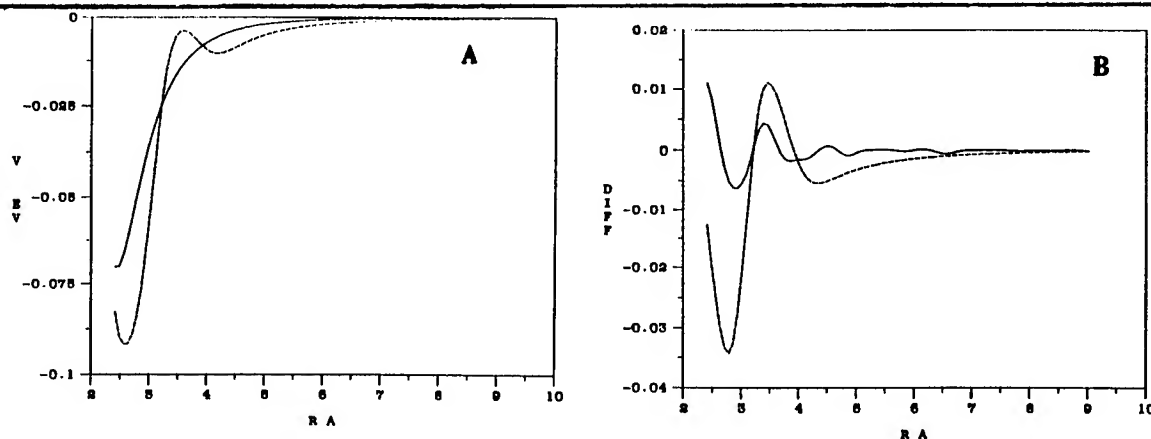


Fig. 8 a) The original Lennard-Jones (LJ) potential (solid line), the final potential which satisfies the conditions discussed in the text and b) the suggested changes (solid line) to the LJ of the maximum entropy calculation and the final changes (dashed line) which produced the dashed potential of a).

available on request). This new potential resulted in a $V(\theta)$ which exactly reproduced the shape of that shown in Fig. 6 but with the minimum at a value of $\theta = 0^\circ$. The fit is so good there is no need to plot the two functions again. This final result could not have been obtained without the very powerful method of "maximum entropy". It is clear that there are countless other situations in which this technique can find applicability.

We are at present investigating the physical significance of the particular shape of this $v(r)$ whose shape is significantly different from that usually used in the description of the H-O interaction. We are

however now permitted to continue with the molecular dynamics simulation of the behavior of nitromethane in the condensed phase confident that the intermolecular potential in hand at least is consistent with most of the measured microscopic quantities. It is hoped that this investigation will shed some light into the problem of the activation energy vs the barrier height. If the present potential is also capable reproducing, for example, the melting temperature then microscopic properties of the liquid should also be amenable to study. It is in the liquid that much work on the detonation properties of nitromethane is available and some theoretical investigations. The latter rely on specific structural mechanisms as explanations of observed phenomena which mechanisms are here accessible for study.

Finally, nitromethane at room temperature freezes at a pressure $\approx .3$ GPa. An X-ray study of the crystal structure as a function of pressure is available to pressures of 6 GPa⁵. In that study is found that the space group is the same, through out this pressure range, as that found above at ambient pressure. Changes occur in the lattice constants and orientation of the molecule but most drastically in the equilibrium orientation of the methyl group. It is stated in that work that the hydrogen atoms could not be located at pressures below 3. Gpa and that when located at higher pressures the methyl group is rotated by 45° with respect to that orientation found at 4 K. Given the pair potential of Fig. 8 and the pressure dependence of the lattice constants, one can calculate $V(\theta)$ as a function of pressure. Fig. 9 presents these results in which $\theta=0^\circ$ corresponds to the equilibrium orientation of the methyl group at 4 K and ambient pressure. Inspection of this figure reveals two significant observations. First, the barrier to reorientation is very shallow at lower pressures but increases rapidly with pressure. This could explain the

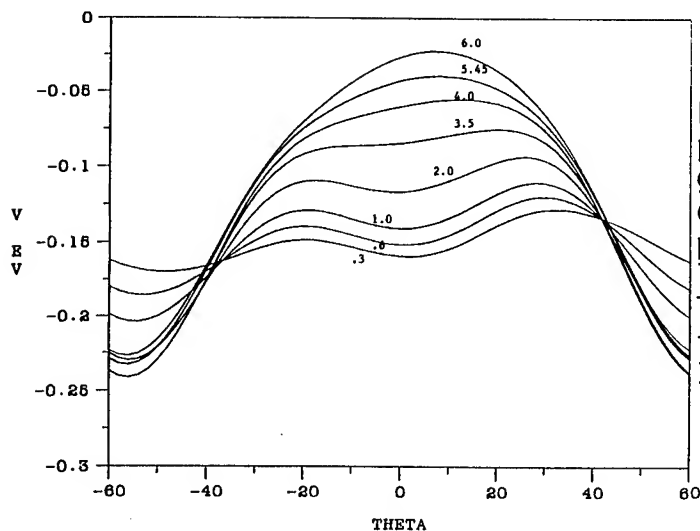


Fig 9. The rotational potential, $V(\theta)$, of the methyl group as a function of pressure obtained from the dashed potential of Fig. 8 and the lattice constants of Ref. 4. The pressures are noted on the figure in units of GPa. (1 GPa = 10 Kbar)

inability to locate the hydrogen atoms at low pressure for they would there be executing essentially free rotations and therefore not be localized for detection. Second, when the barrier is finally large enough to localize the methyl group, its equilibrium position is 50° shifted from 0° . We take these two results as strong confirmation of the utility of this potential. The calculation performed with the original Lennard-Jones potential results in a $V(\theta)$ which approximately obtains six potential minima separated by 60° . Clearly inconsistent with the X-ray pressure study.

References

1. S. F. Trevino, E. Prince and C. R. Hubbard, J. Chem. Phys. **73**, 2996 (1980).
2. S. F. Trevino and W. H. Rymes, J. Chem. Phys. **73**, 3001 (1980).
3. D. Cavagnat, A. Magerl, C. Vettier, I. S. Anderson and S. F. Trevino, Phys. Rev. Lett. **54**, 193 (1985).
4. S. F. Gull and G. J. Danniel, Nature **272**, 686 (1978).
5. D. T. Crommer, R. R. Ryan and D. Schiferl, J. Phys. Chem **89**, 2315 (1985).

Detection of Toxins With A Reversible Biosensor

James J. Valdes

U.S. Army Chemical Research, Development and Engineering Center
Research Directorate, Biotechnology Division
Aberdeen Proving Ground, MD 21010-5423

INTRODUCTION

The production and use of chemical agents in armed conflicts and their potential for use in low-intensity conflicts pose a threat to the security of U.S. forces. The threat is compounded by recent advances in biotechnology, which have blurred the distinction between traditional chemical agents and biological organisms and toxins. Toxins currently used as exotic neurobiological probes in laboratory research may soon be produced in large quantities by advanced fermentation methods and engineered to have precise physicochemical and pharmacologic properties that would make them useful to a military adversary.

The detection of chemicals and toxins in combat, in the environment, or in body fluids has required analytical procedures that, although sensitive, are time-consuming and expensive. Further, these procedures require some knowledge of the identity of the species to be detected. Clearly, a new detection strategy is desirable. The lethal or incapacitating effects of many chemicals and toxins, the pharmacologic effects of drugs, and the action of viruses result from molecular recognition of these agents by physiological targets known generically as receptors. These chemoreceptors are usually located in the cell membrane where they monitor the external environment and respond to the presence of specific chemicals, followed by a cellular response. Receptor-based biosensors are detection devices that utilize a receptor protein as their biological sensing element, and a microsensor as the physical-chemical transducer.

Evanescent fluorosensors can provide a specific and sensitive transducer for a receptor-based biosensor. Although several reports have appeared in the literature using antibody-antigen based evanescent fluorosensors (1, 2), this is the first report where a neurotransmitter receptor protein is coupled to an optic fiber and is used to detect toxins. If successfully used with various receptors, these optic

sensors may become very useful analytical devices for drugs and toxicants that bind to receptors and may develop into significant detection, monitoring, and diagnostic devices. In the case of the fiber-optic system reported herein, the cylindrical waveguide evanescent fluorosensor makes use of the evanescent wave effect by exciting a fluorophore just outside the waveguide boundary (3-5). A portion of the resultant fluorophore emission then becomes trapped in the waveguide and is transmitted back up the fiber. This sensing technique is particularly well suited to a receptor-based biosensor. This is because a fluorescent-tagged ligand bound to a receptor protein, which is immobilized at the fiber surface, can be monitored without interference from the probe in the bulk solution.

The nAChR was the first neurotransmitter receptor to be purified and its molecular properties elucidated (6-9). The receptor genes have been cloned (10, 11) and it is likely to be the first receptor protein that may be produced in large quantities for commercial application. Since the nAChR can easily be harvested in milligram quantities from electric organs of electric fish (12), we immobilized the nAChR (in asolectin lipids) onto the surface of a planar interdigitated capacitive sensor and were able to detect specific concentration-dependent increases in capacitance when an agonist was present (14). This increase was inhibited by antagonists. Although the biosensor could be regenerated, there was great variability in the response depending on the engineering of the capacitance device. Also, several events may have been responsible for the increased capacitance, including binding of agonist to the receptor, which produces changes in receptor conformation, and opening of the receptor's ionic channel, resulting in cation flux through it. The present study was initiated to develop and evaluate another detection device. A fiber optic evanescent fluorosensor and fluoresceinlabelled α -BGT were utilized to provide evidence that the nAChR adsorbed onto the surface of quartz optic fiber could detect ligands that bind to this receptor, whether agonist (e.g. nicotine) or antagonists (e.g. α -BGT or the skeletal relaxant pancuronium). The mechanism of transduction was simple and the change in signal transduction depended only on the receptor-ligand binding step.

MATERIALS AND METHODS

Preparation of Reagents.

α -BGT was obtained from Ventoxin Laboratories (Frederick, MD). Coupling of α -BGT to fluorescein isothiocyanate (FITC) was performed as described by Suszkiw & Ichiki (16) with minor modifications. α -BGT (2 mg) was reacted with 1 mg FITC on celite (Sigma) in 1 ml

of 50 mM bicarbonate buffer at pH 9.5 for 15 min. The celite was removed by centrifugation and the supernatant loaded onto a Sephadex G-25 column (25 x 1.1 cm) and developed with 5 mM ammonium acetate, pH 5.8. The void fractions were pooled, lyophilized, resuspended in 50 mM ammonium acetate, pH 5.8 and loaded onto a CM-52 column (10 x 1.5 cm). The column was developed with the same buffer until the first peak of fluorescence had cleared the column, after which the remaining fluorescent material was eluted with 0.5 M ammonium acetate. Fractions from the second peak were pooled, lyophilized, dissolved in water and assayed for binding to the nAChR. This was detected by its ability to inhibit carbamylcholine-stimulated binding of the allosteric inhibitor [³H]perhydrohistrionicotoxin, as previously reported (17). The FITC labelled α -bungarotoxin (FITC- α -BGT) had an excitation maximum at 493 nm and an emission maximum at 516 nm. FITC- α -BGT at 0.5 μ M in phosphate buffered saline (PBS) (154 mM NaCl in 10 mM sodium phosphate buffer pH 7.4), yielded relative fluorescence (measured on a Gilford Fluoro IV) that was 5 times that of an equimolar solution of FITC in PBS.

Immobilization of Receptor Protein.

Pure nAChR was noncovalently immobilized on the quartz fibers during a 12 hr incubation in a solution which contained 50 μ g/ml nAChR protein in 10 mM sodium phosphate buffer, pH 4.0. Excess receptor was removed from the fiber by one passage through 5 ml of 1 mM sodium phosphate, pH 7.4 or PBS as specified in Results.

Apparatus.

All experimental work was carried out using a fluorimeter designed and built at ORD, Inc., North Salem, NH. The fiber optic evanescent fluorosensor apparatus (Figure 1) was similar in configuration to that reported by Glass, et al. (5). Components of this instrument included a 10W Welch Allyn quartz halogen lamp, a Hamamatsu S-1087 silicon detector, an Ismatec fixed speed peristaltic pump, a Pharmacia strip chart recorder, and bandpass filters and lenses as indicated in the schematic. The quartz fibers, 1 mm in diameter with polished ends, were obtained from ORD, Inc.

The fiber optic evanescent fluorosensor made use of the evanescent wave effect by exciting a fluorophore just outside the waveguide boundary (excitation wavelength - 485/20 nm). A portion of the resultant fluorophore emission then became trapped in the waveguide and was transmitted back up the fiber. This was detected after transmission through 510 nm LP and 530/30 nm filters. The flow cell allowed the center 47 mm of a 60 mm long fiber to be immersed in 46 μ l which was exchanged every 14 sec.

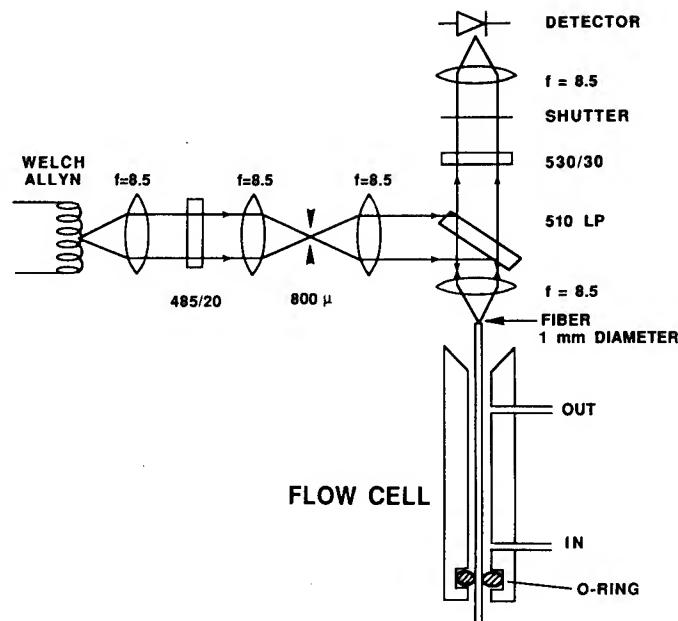


FIG. 1. Schematic presentation of the optical system used to measure fluorescence.

Fluorescence Measurement.

After immobilization of the nACh receptor on the quartz fiber, the fiber was placed in the instrument and perfused for 5 min with PBS containing either bovine serum albumin (BSA) (0.1 mg/ml) alone or BSA containing a receptor ligand where indicated. The fiber was then co-treated with FITC- α -BGT (5 nM) along with the appropriate ligand in PBS containing BSA (0.1 mg/ml). Between experiments, the flow cell was washed with 1% SDS for 2 min followed by PBS for 10 min. Initial rates were determined graphically from tracing of the mV response vs time.

^{125}I - α -BGT Binding.

The specific activity of ^{125}I - α -BGT (100 ci/mmol, from New England Nuclear) was adjusted with unlabelled α -BGT. Quartz fibers with, and without, immobilized nAChR were pretreated for 5 min with PBS containing 0.1 mg/ml BSA. For determination of total α -BGT binding sites, fibers were incubated in the presence of 300 nM ^{125}I - α -BGT for 1 h, washed by 3 passages in PBS, and the radioactivity counted in a LKB Autogamma counter. Inhibition of initial rates of ^{125}I - α -BGT binding by d-tubocurarine (d-TC), were measured using pretreated

(5 min in 0.1 mg/ml BSA in PBS) nAChR-coated fibers. These fibers were incubated for 2 min in 5 nM ^{125}I - α -BGT which bound to quartz fibers lacking immobilized receptors, and was typically 5-10% of the total binding.

RESULTS

The initial rates of binding of FITC- α -BGT to untreated quartz fibers showed a dose-dependent relationship (Figure 2A). These rates were similar for either untreated quartz fibers or quartz fibers coated with immobilized nAChR protein. FITC- α -BGT binding to untreated quartz fibers, however, was partially inhibited (at 500 nM FITC- α -BGT) or completely inhibited (at 5 nM FITC- α -BGT) by pretreatment of the fibers with BSA (0.1 mg/ml) (Figure 2; Table 1). Furthermore, pretreatment of the nAChR-coated fibers for 30 min with 5 nM unlabelled α -BGT resulted in 75% inhibition of the initial rate of FITC- α -BGT binding (Table 1).

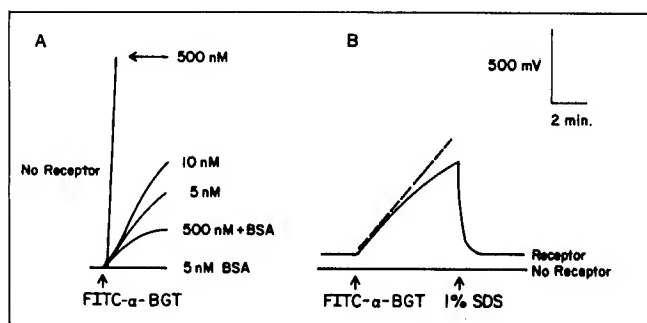


FIG. 2. Binding of fluorescently labeled α -BGT to (A) untreated quartz fibers in the presence or absence of BSA (0.1 mg/ml in PBS buffer), (B) untreated or nAChR-coated quartz fibers. FITC- α -BGT was introduced at 5 nM in PBS containing BSA (0.1 mg/ml). Dashed line represents graphically determined initial rates. Receptor and FITC- α -BGT were washed from the fiber with 1% sodium dodecyl sulfate (SDS).

TABLE 1
Effect of BSA Pretreatment on Nonspecific Binding
of FITC- α -BGT

Fiber type ^a	Treatment ^b	Initial rate ^c (mV/min)
Quartz untreated	FITC- α -BGT	108
Quartz untreated	FITC- α -BGT, BSA	0
Quartz with nAChR	FITC- α -BGT	83
Quartz with nAChR	FITC- α -BGT, BSA	107
Quartz with nAChR	FITC- α -BGT, BSA (pretreated with α -BGT)	25

^a Untreated fibers were washed with methanol prior to use. Quartz fibers were coated with nAChR as described under Materials and Methods.

^b Fibers were perfused at 0.2 ml/min with PBS buffer containing FITC- α -BGT (5 nM) or BSA (0.1 mg/ml) or pretreated with unlabeled α -BGT (5 nM, 30 min), as specified. Fibers treated with BSA were pretreated for 5 min prior to introduction of FITC- α -BGT.

^c Initial rates were determined graphically from the first 30 s of data. Rates are representative of at least three experiments.

These results suggested that when nonspecific FITC- α -BGT binding was eliminated by treatment of the fibers with BSA, the specific binding of FITC- α -BGT to the nAChRs (noncovalently immobilized on these fibers) was responsible for the observed fluorescence. The fluorescence response of the receptor-coated fibers increased for 30 min then remained stable for several hours.

The nAChR optic fibers bound ^{125}I - α -BGT in a dose-dependent manner (Figure 3A). Increasing ratios of ^{125}I - α -BGT to unlabelled α -BGT resulted in a linear increase in the amount of ^{125}I - α -BGT bound/fiber. The receptor-coated fibers bound 4.2 ± 0.4 pmol ^{125}I - α -BGT sites/fiber ($n = 9$). Increasing the ratios of FITC- α -BGT to nonlabelled α -BGT, at a total concentration of 3 μM , also yielded a linear increase in the maximum fluorescence response (Figure 3B). The fluorescence response for the nAChR-coated fibers saturated after 30 min and yielded a maximum value of 8.4 V/fiber for FITC- α -BGT. The noise level of the fluorescence response during a typical experiment was in the range of 10-50 mV. This translates into a maximum sensitivity of 25 fmol of nAChR which could be detected on a fiber, and a signal to noise ratio of >99:1.

The initial rate of FITC- α -BGT (5 nM) binding to the receptor-coated fibers was inhibited in a dose-dependent manner by the competitive antagonist d-tubocurarine (d-TC) (Figure 4; Figure 5). However,

about 20% of the maximum binding remained even at 1 mM d-TC. The initial rate of ^{125}I - α -BGT (5 nM) binding to the nAChR-coated fibers was also inhibited by d-TC (Figure 5). The displacement curves were similar, though that for FITC- α -BGT was shifted to the right. This was presumably due to the effect of the residual rate of FITC- α -BGT binding even at high concentrations of d-TC.

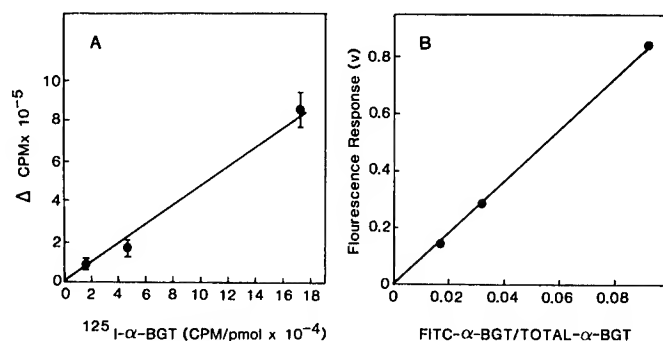


FIG. 3. Specific binding of α -BGT to nAChR-coated optic fiber. (A) Binding of α - ^{125}I -BGT to fibers as a function of the specific activity of α - ^{125}I -BGT. Total α -BGT concentration was 300 nM. Symbols and bars represent means of triplicate determinations \pm SEM. (B) Maximum fluorescence response as a function of the ratio of FITC- α -BGT to nonlabeled α -BGT. Total α -BGT concentration was 3 μM . Data are representative of at least two measurements.

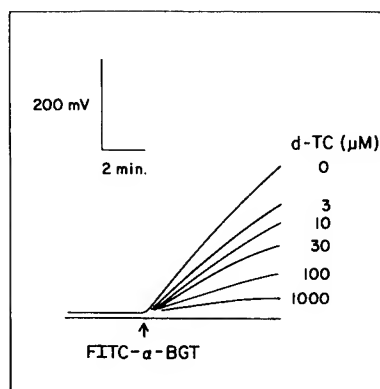


FIG. 4. The concentration-dependent inhibition of receptor-mediated binding of FITC- α -BGT (5 nM) by the competitive antagonist d-TC.

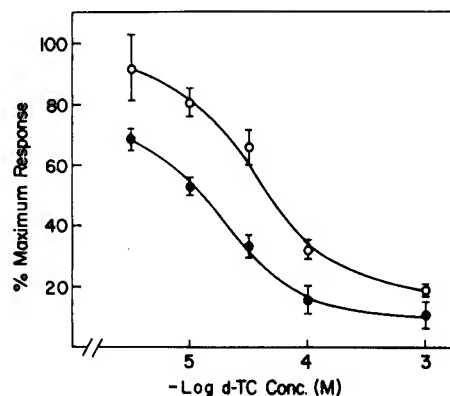


FIG. 5. The effect of various concentrations of *d*-TC on the initial rates of binding of 5 nM of α -¹²⁵I-BGT (●) and FITC- α -BGT (○) to nAChR-coated fibers. In both cases, the fibers were coincubated with *d*-TC or the labeled α -BGT. Symbols and bars are means of triplicate experiments and represent \pm SEM.

α -BGT, *d*-TC and the agonist carbamylcholine, all of which bind to the same receptor site, inhibited binding of FITC- α -BGT to the nAChR-coated fibers, and resulted in a dose-dependent decrease in fluorescence (Figure 6).

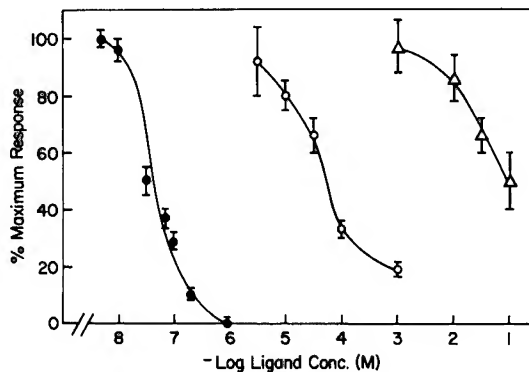


FIG. 6. The effects of various concentrations of *d*-TC (○), carbamylcholine (Δ), and α -BGT (●) on binding of 5 nM FITC- α -BGT to nAChR immobilized on quartz fibers. The various ligands, at the appropriate concentrations, were coperfused with FITC- α -BGT. There was no premeasurement treatment of the fibers with ligands/BSA alone. Symbols and bars are means of triplicate measurements \pm SEM.

The relative potencies of d-TC, carbamylcholine and unlabelled α -BGT in inhibiting the initial rates of binding of FITC- α -BGT (5 nM) to the fiber-coated nAChR, were similar to those reported for detergent extracts of the purified nAChR from Torpedo (18, 19). However, the IC₅₀ (concentrations giving 50% inhibition) values for d-TC and carbamylcholine were shifted in order of magnitude to the right of that expected for detergent extracts of purified nAChR. Furthermore, when the receptor-coated fibers were pretreated with 1 mM carbamylcholine for 10 min prior to measuring FITC- α -BGT displacement, it gave 42% inhibition (Table 3) instead of only 10% (Figure 6) when the fibers were not pretreated.

The initial rates of FITC- α -BGT binding to nAChR were not affected by concentrations of NaCl up to 154 mM (Table 2).

TABLE 2
Effect of Ionic Strength on FITC- α -BGT Binding

Na phosphate buffer ^a (mM)	NaCl (mM)	d-TC (μ M)	Initial rate ^b (mV/min)
10	154	0	103 \pm 9
10	154	30	60 \pm 10
1	0	0	113 \pm 7
1	0	30	50 \pm 5

^a Buffer contained 0.1 mg/ml BSA. Fibers were pretreated for 5 min with the buffer indicated before introduction of 5 nM FITC- α -BGT in the appropriate buffer in the absence or presence of d-TC.

^b Rates \pm SEM ($n = 3$).

Similarly, inhibition of FITC- α -BGT binding by d-TC was also unaffected by ionic strength. On the other hand, the initial rates of FITC- α -BGT binding to nAChR-coated fibers were sensitive to cholinergic ligands (Table 3). Several receptor agonists and antagonists effectively blocked the binding of FITC- α -BGT and reduced the rate of increase in observed fluorescence.

TABLE 3
Inhibition by Receptor Ligands of the Binding of FITC- α -BGT to the nAChR on the Optic Fibers

Drug ^a	% inhibition of initial rate of fluorescence change ^b
Agonists	
Acetylcholine	61
Carbamylcholine	42
Nicotine	27
Depolarizing blockers	
Decamethonium	52
Succinylcholine	45
Competitive antagonists	
<i>d</i> -Tubocurarine	87
Pancuronium	76
<i>Naja</i> α -neurotoxin (10 μ M)	100

^a Unless otherwise specified, the drug was present at 1 mM in PBS containing BSA (0.1 mg/ml) and FITC- α -BGT (5 nM). Fibers were pretreated with the indicated drug for 10 min prior to addition of the FITC- α -BGT.

^b Initial rate of fluorescence change in the absence of inhibitor = 100%. The values are means of two experiments which varied by <10%.

DISCUSSION

The intensity of the observed fluorescence is dependent on the number of fluorescent molecules within the boundaries of the evanescent wave. FITC at 5 nM in PBS, is not detected by the fluorosensor, apparently because fluorescein does not bind well to quartz glass. However, 1 nM FITC- α -BGT, a concentration which gives equivalent fluorescence to 5 nM FITC solutions, produces a good fluorescence signal from the evanescent fluorosensor. Using the slope of the initial change in fluorescence as a parameter, it is evident that FITC- α -BGT binds to the quartz fiber in a dose-dependent manner (Figure 2). In the presence of 0.1 mg/ml BSA, this nonspecific binding of 5 nM FITC- α -BGT to the quartz fiber is totally eliminated (Figure 2A, B). This is a concentration of BSA that does not interfere with binding of FITC- α -BGT to the nAChR (Figure 2B, Table 1). Accordingly, it may be concluded that the observed fluorescence in the BSA-treated fibers is a result of FITC- α -BGT binding to the nAChR and not the result of its nonspecific binding to the quartz fiber.

The binding of α -BGT to the nAChR is quasi-irreversible event, making it difficult to study competition displacement of α -BGT binding by reversibly binding drugs under equilibrium conditions. Therefore, affinities of receptor agonists and antagonists have been measured accurately by monitoring their effects on the association rate of α -BGT at pre-equilibrium conditions (18, 19). These reversible ligands, at physiological concentrations, have been shown to decrease the rate of toxin binding in a dose-dependent manner.

This optical biosensor may also be used to quantitate protein in solutions. The amount of nAChR picked by the optic fiber, determined from ^{125}I - α -BGT binding experiments (Figure 3A), is 4.2 pmol/fiber or 29 fmol/mm². This value is similar to that reported by Bhatin et al. (20) for covalent immobilization of IgG to silica fibers (i.e. 6 fmol/mm²). The maximum fluorescence given when all receptor sites bind FITC- α -BGT is 8.4 V/fiber (Figure 3B). Considering the very small noise level of less than 1% of the maximum signal (i.e. 10-50 mV), as little as 25 fmol of receptor per fiber are sufficient to give a significant signal.

The finding that binding of FITC- α -BGT to the immobilized nAChR on the optic fiber is affected by nicotinic cholinergic ligands (Table 3, Figures 4-6), indicates that the recognition function of the nAChR protein is intact. Consequently, this technique is applicable for detection of any cholinergic ligand which competes for the α -BGT binding site. The IC₅₀ for d-TC and carbamylcholine are about an order of magnitude higher on binding of FITC- α -BGT, compared to their IC₅₀ values on binding of ^{125}I - α -BGT to detergent extracts of Torpedo nAChRs (21) (i.e. d-TC 6.4 μM , carbamylcholine 50 μM). These results suggest a change in the fiber-immobilized receptor, which may be due to the fact that the nAChR fluorosensor utilizes pure receptor protein in absence of detergents or lipids. It may be possible to enhance the sensitivity of the sensor to cholinergic ligands by severalfold if the environment of the nAChR on the fiber is lightly modified, such as by addition of lipids after receptor immobilization on the optic fiber.

The optic fiber nAChR biosensor shows high sensitivity towards α -BGT, detecting concentrations less than 1 μM , compared to >50 nM for another biosensor using an electrode transducer (15). Another important advantage is that the mechanism underlying signal production in this biosensor is clearly understood, while the mechanism is unknown for three others: the ion-sensitive field-effect transistor (13), the capacitive sensor (14) and the interdigitated electrode (15). Furthermore, the optic-fiber biosensor has high signal to noise ratio of 99:1. By contrast, the ion-sensitive field-effect transistor (13) and the interdigitated electrode (15) nAChR biosensors have

a signal to noise ratio of only 0.1-0.2 and could operate only in very low ionic-strength media. The insensitivity of the fluorescent signal to high salt concentration is an important feature that would allow the biosensor to operate well in high ionic-strength such as found in either in body fluids (e.g. plasma and urine) or the environment. Another advantage of the nAChR optic fiber fluorosensor is that binding of α -BGT is a more receptor-specific signal than a response to acetylcholine in biosensor assays especially when the mechanism of transduction is still unknown. Muscarine receptor and acetylcholinesterase are two proteins that would react with acetylcholine and interfere with signal production in capacitive or electrode sensors. In an earlier study, this factor was addressed by comparing pure and partially purified nAChR preparations (14).

A disadvantage of using an irreversible FITC-labelled ligand is that the biosensor cannot be regenerated and reused. Although a disposable sensor may be useful, a reversible fluorescent probe would allow multiple uses of the single fiber. Use of GI conotoxin as a receptor probe has recently been shown by Eldefrawi, Rogers and Valdes (data not shown) to provide a reversible monitor of toxin-receptor interactions with an "all clear" signal. This could be monitored for hours. This technology has extraordinary applications in the areas of environmental monitoring, medical diagnostics, and industrial process control, as well as for the detection of CBW agents, explosives and drugs.

REFERENCES

1. Tromberg, B.J., Sepaniak, M.J., Vo-Dinh, T. and Griffin, G.D. (1987) Anal. Chem. 59, 1226-1230.
2. Anderson, F.P. and Miller, G.W. (1988) Clin. Chem. 34, 1417-1421.
3. Hirschfeld, T.B. (1965) J. Can. Spectrosc. Nov. 126.
4. Hirschfeld, T.B. and Block, M.J. (1984) "Fluorescent Immunoassay Employing Optical Fiber in Capillary Tube," U.S. Patent 4,447,546.
5. Glass, T.R., Lackie, S. and Hirschfeld, T. (1987) Applied Optics 26, 2181-2187.
6. Olsen, R.W., Meunier, J.C. and Changeux, J.-P. (1972) FEBS Letters 28, 96-100.
7. Schmidt, J. and Raftery, M.A. (1973) Biochemistry 12, 852-856.

8. Eldefrawi, M.E. and Eldefrawi, A.T. (1973) Arch. Biochem. Biophys. 159, 362-373.
9. McNamee, M.G., Weill, C.C. and Karlin, A. (1975) Ann. N.Y. Acad. Sci. 264, 175-182.
10. Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Fusutani, Y., Kikuyotani, S., Hirose, T., Asai, M., Takashima, H., Inayanura, S., Miyata, T. and Numa S. (1983) Nature 301, 251-255.
11. Claudio, T., Ballivet, M., Patrick, J. and Heinemann, S. (1983) Proc. Natl. Acad. Sci. USA 80, 1111-1115.
12. O'Brien, R.D., Eldefrawi, M.E. and Eldefrawi, A.T. (1972) Ann. Rev. Pharmacol. 12, 19-34.
13. Gotoh, M., Tarniya, E., Momoi, M., Kagawa, Y. and Karube, I. (1987) Analyt. Lett. 20, 857-870.
14. Eldefrawi, M.E., Sherby, S.M., Andreou, A., Mansour, N.A., Annau, Z., Blum, N.A. and Valdes, J.J. (1988) Analyt. Lett. 21, 1665-1680.
15. Taylor, R.F., Marenchic, I.G. and Cook, E.J. (1988) Analyt. Chim. Acta. 213, 131-138.
16. Suszkiw, J.B. and Ichiki, M. (1976) Anal. Biochem. 73, 109-114.
17. Sherby, S.M., Eldefrawi, A.T., Albuquerque, E.X. and Eldefrawi, M.E. (1985) Molec. Pharmacol. 27, 343-348.
18. Weber, M., David-Pfeuty, T. and Changeux, J.-P. (1975) Proc. Natl. Acad. Sci. USA 72, 3443-3447.
19. Blanchard, S.G., Quast, U., Reed, K., Lee, T., Schimerlik, R.V., Claudio, T., Strader, C.D., Moore, H.H. and Raftery, M.A. (1979) Biochemistry 18, 1875-1883.
20. Bhatia, S.K., Shriver-Lake, L.C., Prior, K.J., Georger, J.H., Calvert, J.M., Bredehorst, R. and Ligler, F.S. (1989) Anal. Biochem. 178, 408-413.
21. Schmidt, J. and Raftery, M.A. (1974) J. Neurochem. 28, 617.
22. Rechnitz, G.A. (1988) Chem. Engineering News Sep. 5, 24-36.

Simultaneous Determination of Temperatures and OH
Concentrations in a Solid Propellant Flame

*John A. Vanderhoff, Dr.

Anthony J. Kotlar, Dr.

US Army Ballistic Research Laboratory
Aberdeen Proving Ground, MD 21005-5066

I. INTRODUCTION

Understanding a combustion event usually requires knowledge of the temperatures, species involved and their concentration profiles. Thermochemical equilibrium calculations will produce final product species identity and concentration as well as adiabatic flame temperatures; however these results are not very event specific. Information about the detailed chemical pathways from reactants to final products comes from studying the chemically active transient species. These species are generally present in small quantities (<1%) and exist primarily in a zone of small extent: the reaction zone.

The small concentrations of transient combustion species (radicals) renders the use of Raman techniques unattractive for concentration measurements. Non-linear techniques are not considered since they are not simply related to concentration. Laser induced fluorescence (LIF) and absorption have sufficient sensitivity for the detection of the radical species. However, although LIF is a linear technique, it also is not simply related to concentration because of the quenching effects on the excited state which are exacerbated with increasing pressure. Absorption is directly related to concentration and can provide useful information under conditions where a line-of-sight measurement is meaningful.

This paper has a two-fold objective. First, an absorption technique is described which can simultaneously measure temperature and concentration profiles in time varying hostile combustion environments. Second, the viability of this technique is demonstrated by obtaining temperature and OH concentration profiles from a solid propellant flame burning at elevated pressure. This absorption technique is a variation of the method described by Lempert¹. The present technique, however, has a different light source and detector which allows the OH absorption spectrum to be obtained and analyzed in much more detail as well as allowing

different species to be studied with the same experimental setup. A description of the absorption experiment and data analysis is given in the next two sections of the paper.

II. EXPERIMENTAL

Many investigators have determined OH concentrations from the absorptions in the 306.4 nm system using lamp light sources and a photomultiplier detector. The present arrangement incorporates a high intensity mercury-xenon arc lamp light source but the detector is an intensified photodiode array which can be much more suitable for probing transient combustion processes. As an indication of the sensitivity of the experimental technique, absorptions of 0.15% were readily detectable in a steady state flame environment. The optical path for the absorption measurements of propellant burning within a pressurized combustion vessel is shown on Fig. 1. Two 20 cm focal length convex lenses (L2) focus and recollimate the light beam from the 1 kw mercury-xenon arc lamp and two apertures and a 0.15 mm pinhole confine the spatial extent to approximately 0.15 mm in the burning direction. The aperture between the sample and the spectrometer minimizes collection of emission signal and the 10 cm focal length cylindrical lens (CYL1) focusses the transmitted light onto the spectrometer entrance slit. Aperture sizes (A) between 0.5 and 1.0 mm were used to keep the emission signal levels to less than 2% of the absorption signal. In actuality the aperture size was set from the results obtained from a $\text{CH}_4/\text{N}_2\text{O}$ flame, and it was assumed that the propellant would be similar in this respect. The spectrometer has a 0.3 m focal length and a 2400 groove per mm grating which when operated second order gave a spectral resolution of about 0.03 nm with a 0.025 mm entrance slit. This system allows a 6 nm band of light to be sampled at one time on the intensified photodiode array. A uv transmitting black glass filter (Schott UG-11) eliminated first order light from entering the spectrometer. OH rotational absorptions of interest are much narrower than this 6 nm band; thus intensities of wavelengths which undergo absorption, and nearby wavelengths which do not, can be recorded simultaneously. The system therefore provides a simple calibration of effective incident intensity in cases where continuous absorbers or scatterers interfere.

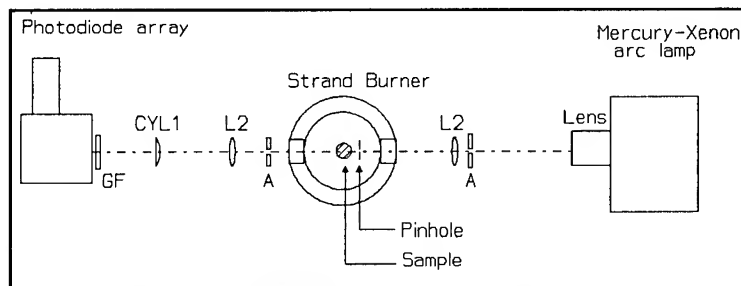


Figure 1 The light path for the optical absorption experiments. The glass filter (GF) was used to eliminate first order light when operating the spectrometer second order; the sample is either a premixed flame or propellant.

Two combustion systems have been experimentally studied: a premixed $\text{CH}_4/\text{N}_2\text{O}$ flame and a solid propellant flame. The solid propellant (HMX1)² is composed of 73% cyclotetramethylenetetranitramine (HMX). The $\text{CH}_4/\text{N}_2\text{O}$ flame has been used primarily for

calibration and comparison purposes where a path length has been chosen similar to the path length used for the propellant absorption studies. A burner consisting of many small holes within a 0.4 cm diameter (the path length) was used to support this premixed flame. The top of the combustion vessel is removable and thus this flame can be placed at the sample location for experimentation. Long path lengths are traditionally desirable for trace species since, for small absorptions, the absorption increases linearly with path length. However, in these studies the overriding concern was to produce a horizontal burning propellant surface, *i.e.*, a surface which is parallel to the direction of the arc lamp light beam. This condition could only be obtained for small cross sections. An HMX1 propellant geometry which exhibited a reasonable cigarette fashion burn and was thus used, was a strand about 4 cm long with an octagonal cross section that measured 0.65 cm between opposing faces. The propellant strand was coated with a thin layer of fingernail polish to inhibit side burning.

A partial diagram of the windowed combustion vessel is shown in cross section on Fig. 1. This vessel is made from stainless steel with two pairs of opposing windows and provides an environment to burn propellants at an elevated pressure. An exit orifice in the vessel allows a purge gas to be flowed at rates typically about four times the propellant gasification rate. A video camera (not shown) provides a photographic record of each propellant burn. This record provides the burning rate of the propellant as well as visual information on the "flatness" of burn. Further details of the combustion vessel can be found elsewhere.^{3,4}

Propellant burning in this experiment is a transient process since the burning surface regresses with time. This limits the amount of time that data can be taken depending on the amount of spatial resolution desired. The first studies of OH in a propellant produced marginally acceptable data and thus a propellant feed mechanism was incorporated into the experiment to expand the data acquisition time. A sketch of the feed mechanism is shown on Fig. 2. This mechanism is more rudimentary than previous feed mechanisms^{5,6} in several respects. First, the feed mechanism is not enclosed in the

pressure vessel, but rather couples in linear motion by a drive shaft sealed with a plastic ferrule swage. Second, the drive shaft is unidirectional during the experiment with the feed rate and travel distance having been programmed into the speed control. Instead of maintaining an optical feedback control, the feed rate is set to be some large fraction of the propellant burn rate. For the experiments reported here the speed controller was programmed to drive the translator at a linear rate of 1.00 mm/s. The apparent burn rate, as viewed by

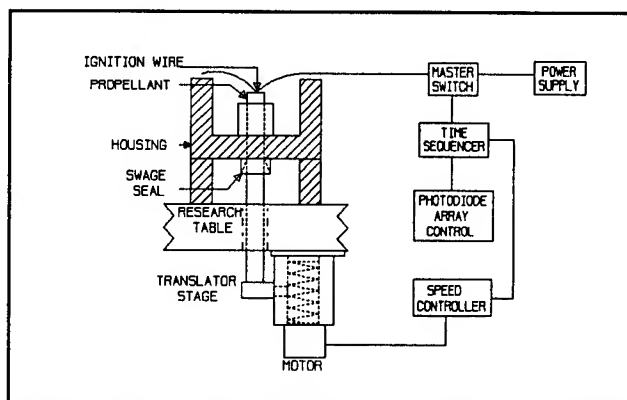


Figure 2 A sketch of the major components of the propellant feed mechanism and associated components that are triggered at various times during the course of an experimental run.

a video camera, was measured as 0.27 mm/s which means the actual burn rate of HMX1 in a combustion vessel pressurized to 1.5 MPa nitrogen is 1.27 mm/s.

The following sequence of events occur for a typical propellant burn experiment. The combustion vessel is pressurized to 1.5 MPa nitrogen and the video recording system is activated. A master switch engages the power to heat the ignition wire and also starts the time sequencer. After preset delays the time sequencer starts the propellant feed and triggers the photodiode array control. The apertured light from the arc lamp is terminated by the propellant strand at the beginning of each experiment and the initial passage of this light to the detector indicates when the propellant surface is in the sampling region. The photodiode array repetitively scans and resets many times while storing each scan into buffer memory. The total data accumulation time as well as the time for each scan can be varied over a wide range. Total data accumulation times were major fractions of the total propellant burn time, ranging from 15 to 20 seconds.

III. DATA ANALYSIS

In this section working equations are presented which are used to extract temperatures and absolute concentrations from experimental absorption spectra. The development and detailed discussions of these equations can be found elsewhere.^{7,8} Based on the differential absorption law for a parallel beam of light of frequency ν traveling in the $+x$ direction through a medium with absorption coefficient $k(\nu)$,

$$-dI = I k(\nu) dx ,$$

the intensity of the incident beam, I_0 (assumed to be constant), is attenuated along a path of length ℓ according to

$$I(\nu) = I_0 \exp[-\ell k(\nu)] .$$

Assuming a Boltzmann distribution the transmitted intensity of a group of lines is given by

$$I(\nu) = I_0 \exp\{ -(h\nu\ell/c) [N_T/Q(T)] [\sum_j B_j g_j \exp(-E_j/kT) P_j(\nu)] \} ,$$

where h is Planck's constant, c is the speed of light, k is the Boltzmann constant, T is the absolute temperature, B_j is the absolute Einstein coefficient of absorption for the j th transition in energy density units, g_j is the degeneracy of the j th sublevel and E_j its energy. N_T is the total number density, $Q(T)$ is the partition function which is evaluated using a standard closed form expression, and P_j is the transition lineshape. For conditions where the light source and spectrometer bandwidth is much larger than the width of a typical absorption line, an instrument function, $S(\nu, \nu_0)$, centered at ν_0 and normalized to unit area is introduced to give

$$I_{tr} = \int S(\nu, \nu_0) I(\nu) d\nu ,$$

where I_{tr} is the integrated light transmitted.

The Einstein coefficients for OH are from Dimpfl and Kinsey⁹ and the transition frequencies are from Dieke and Crosswhite¹⁰. 126 transitions of the A-X (0,0) band are included in the calculation, although typically only 90 are used for the range of data selected for analysis. Other molecule specific parameters are from Huber and Herzberg¹¹. A multi-parameter, nonlinear, least squares program with interactive graphics running on a 386/25MHz PC was used to fit the data. Since the width of the instrument function, taken to be Lorentzian, is larger than the transition linewidth, and in order to reduce computation time, the transition lineshape was approximated as a delta-function, thereby reducing the above integral to a sum. Using this model, the time to fit a spectrum was 10 minutes or less. Further details about the fitting procedure are found in the Appendix.

IV. RESULTS

Previously⁴, OH concentrations for an HMX propellant have been computed from peak absorptions of the Q₂ bandhead in the A-X (0,0) vibrational band system for OH. There the experimental data were not sufficiently resolved for the determination of temperature and thus it was necessary to construct a temperature profile by other means⁴. Only for a CH₄/N₂O flame were the data of sufficient quality to calculate a temperature from a Boltzmann plot⁴. We report here simultaneous temperature and concentration determinations for both a steady state CH₄/N₂O flame and a solid propellant flame. An example OH absorption spectrum, taken in a CH₄/N₂O flame at atmospheric pressure, is shown in Fig. 3. The absorption path length for this flame was 0.4 cm and one second was required to obtain the data. All of the absorptions observed experimentally on Fig. 3 are accounted for by the fitting program. Thus there is reasonable confidence that the discrete absorptions over the wavelength range from 307.7 to 310.1 nm are

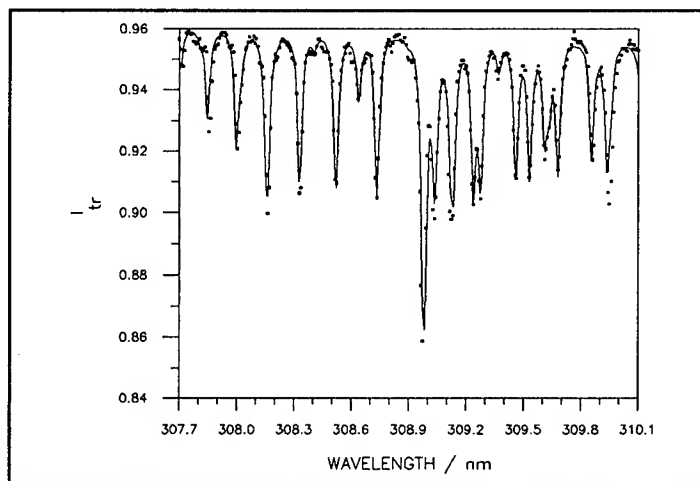


Figure 3 Absorption spectrum of the A-X (0,0) band system of OH, taken 1 mm above the burner surface of a premixed CH₄/N₂O flame. (■) is the data and the solid line is the least squares fit.

due solely to OH. A value of 2367 K has been determined for the temperature of this small premixed flame. All least squares fitted values and the standard deviations are given in Table 1. This value is in good agreement with prior determinations⁴ of temperature for this flame which gave 2360 K and 2304 K obtained from Boltzmann plots of peak rotational values for OH emission and absorption, respectively. The adiabatic flame temperature for a stoichiometric $\text{CH}_4/\text{N}_2\text{O}$ flame is 2922 K; however, this flame was strongly burner stabilized and, as has been found with other burner stabilized flames¹², can have substantial heat loss to the burner head. Although the purpose of using a small steady state flame was only for setting up for propellant measurements, the absolute concentration determination for OH agrees well with published trends. That is, the fitted value of 3.9×10^{16} molecules/cc is about 1.7 times the thermochemical equilibrium value¹³ and is very similar (1.5 times) to the peak concentration of OH¹² reported for a stoichiometric $\text{CH}_4/\text{N}_2\text{O}$ flame operating at atmospheric pressure.

Sample OH absorption data together with the least squares fit for an HMX1 propellant sample burning in 1.5 MPa nitrogen are shown on Fig. 4. The fitted values for the spectrum of Fig. 4 and for eight other spectra at different positions from the propellant surface are tabulated in Table 1. It was necessary to accumulate data for times corresponding to movement of the propellant surface in order to obtain statistically reasonable data; hence the propellant spectrum of Fig. 4 is an average over 0.4 mm. Here again for the propellant case all of the prominent features over the wavelength range from 307.9 to 310.3 nm have been accounted for with the OH fitting program. Temperatures and OH concentrations with their standard deviations ($\pm 1\sigma$) have been tabulated in Table 1 and are plotted on Figs. 5 and 6. The temperature rise from the propellant surface to the final flame temperature occurs within 0.4 mm of the propellant surface. Within the statistical uncertainty of the data the fitted temperature reaches the adiabatic flame temperature. It also appears from Fig. 5 that the temperature is slightly dropping with increasing distance from the propellant surface. These temperature measurements can be directly compared with the recent results obtained by Stufflebeam and Eckbreth¹⁴. Using a CARS technique they have measured temperatures in HMX1 propellants burning in 2.31 MPa

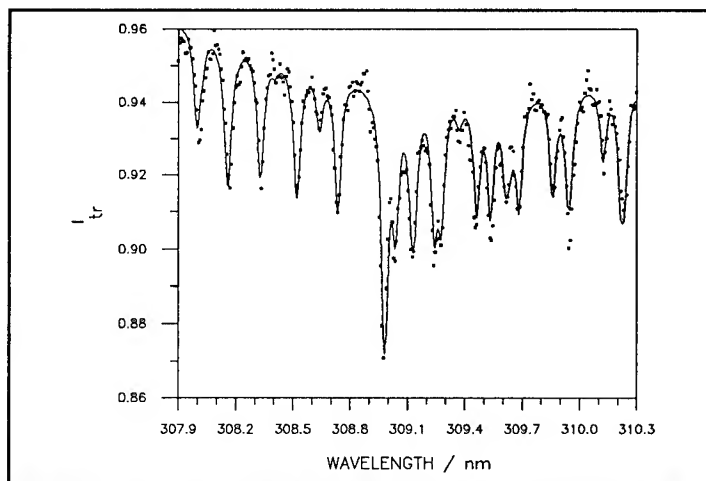


Figure 4 Same as Fig. 3 except here the sample is an HMX1 propellant burning in 1.5 MPa N_2 from 0.4 to 0.9 mm above the propellant surface, corresponding to a data acquisition time of 1.6 seconds.

helium. Values ranging from 2600 K at 2.4 mm from the propellant surface to 1900 K at 12 mm from the surface were observed. The present data are in excellent agreement with these published results.

OH concentrations as a function of distance from the propellant surface are shown on Fig. 6. A rise in OH concentration can be observed at distances less than about 1 mm from the propellant surface and the concentration decreases further from the propellant surface due to cooling of the flame. Together with these data are dashed lines which represent calculated OH con-

Table 1 Temperature and OH concentration values and standard deviations obtained from fitting OH absorption spectra.

System	Distance	Temperature	OH Concentration
	mm	K	10^{16} molecules/cc
CH / N O 4 2	1	2356(124)	3.9(0.2)
HMX1	0 - 0.4	2739(281)	1.5(0.2)
	0.4 - 0.9	2512(115)	2.5(0.1)
	0.9 - 1.8	2690(116)	2.8(0.2)
	1.8 - 2.7	2540(103)	2.7(0.1)
	2.7 - 3.6	2479(114)	2.2(0.1)
	3.6 - 4.5	2704(131)	2.2(0.1)
	4.5 - 5.4	2038(103)	1.6(0.1)
	5.4 - 6.3	2290(107)	1.7(0.1)
	6.3 - 7.2	2437(176)	1.2(0.1)

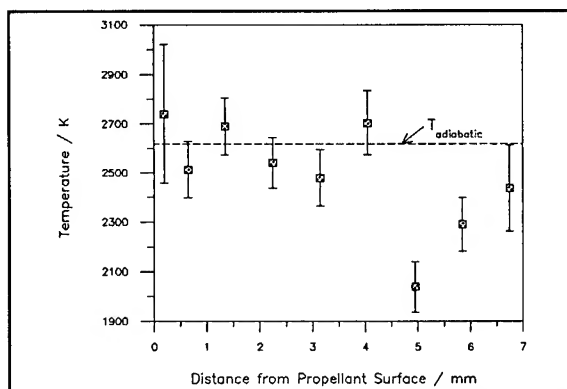


Figure 5 Temperature as a function of distance from the propellant surface for an HMX1 propellant burning in 1.5 MPa nitrogen. (■) represent the temperature data which is tabulated in Table 1.

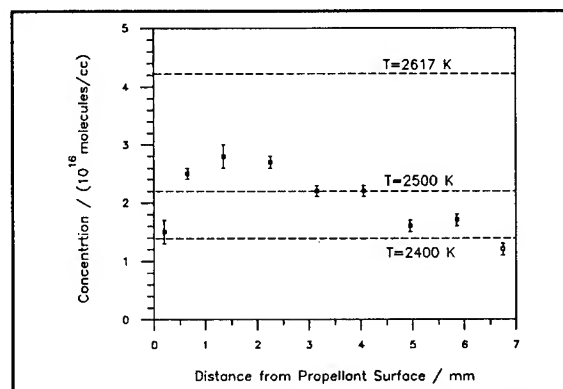


Figure 6 Concentration as a function of distance from the propellant surface corresponding to the temperatures in Fig. 5.

centrations assuming thermochemical equilibrium. The three concentrations computed are for temperatures of 2617 (adiabatic), 2500 and 2400 K. Thus it is seen from Fig. 6 that the OH concentrations computed from fitting OH absorption spectra fall in a thermochemical equilibrium concentration range from slightly above 2500 K to about 2400 K. Although the standard deviation in the OH concentration determinations do not allow for concentrations high enough to satisfy an adiabatic flame temperature, the total error in the OH concentration determination is estimated to be a factor of two. This additional uncertainty comes from the path length and lineshape variables which affect the concentration determination much more than the temperature. A factor of two uncertainty in experimental OH concentration does include the computed OH equilibrium concentrations appropriate for the adiabatic flame temperature. There is no known direct comparisons for absolute OH concentrations measured in a solid propellant flame except for some of our previous work⁴ using only a peak absorption of the Q₂ bandhead at 309 nm. Here concentrations for OH in HMX1 burning in 1.5 MPa nitrogen were about a factor of two smaller than those presented here. The difference in these results is primarily the result of using a Lorentzian instrument function for the present study, rather than the triangular function previously used.

V. SUMMARY

Temperature and OH concentration profiles have been obtained for a solid propellant burning at elevated pressure using a sensitive absorption technique together with least squares fitting of a number of rotationally resolved lines. The rather coarse spatial resolution of 0.4 mm compromised the ability to observe the initial rise in temperature and OH concentration however, because OH varies slowly in comparison to many other radical species (after this initial rise), it is believed that the rest of the profile is adequately represented. Both the temperature and concentrations obtained are consistent with this solid propellant flame burning at its adiabatic flame temperature, and an OH concentration in thermochemical equilibrium.

ACKNOWLEDGEMENT

We thank Dr. Tim Edwards of AFAL for supplying the HMX1 propellant used in this study.

REFERENCES

1. Lempert, W. R.: *Microwave Resonance Lamp Absorption Technique for Measuring Temperature and OH Number Density in Combustion Environments*, Combust. Flame, **73**, 89 (1988).
2. Edwards, J. T.: *Solid Propellant Flame Spectroscopy*, AFAL-TR-88-076, 1988.
3. Vanderhoff, J. A.: *Spectral Studies of Propellant Combustion: Experimental Details and Emission Results for M-30 Propellant*, BRL-MR-3714, 1988.
4. Vanderhoff, J. A.: *Spectral Studies of Solid Propellant Combustion II. Emission and Absorption Results for M-30 and HMX1 Propellants*, BRL-TR-3055, 1989.
5. Rekers, R. G. and Villars, D. S.: *Flame Zone Spectroscopy of Solid Propellants*, Rev. Sci. Instrum., **25**, 424 (1954).
6. Edwards, T. J., Weaver, D. P., Adams, R., Hulsizer, S., and Campbell, D. H.: *High-Pressure Combustor for the Spectroscopic Study of Solid Propellant Combustion Chemistry*, Rev. Sci. Instrum., **56**, 2131 (1985).
7. Mitchell, A. C. G. and Zemansky, M. W.: *Resonance Radiation and Excited Atoms*, Cambridge University Press, 1971.
8. Lucht, R. P., Peterson, R. C. and Laurendeau, N. M.: *Fundamentals of Absorption Spectroscopy for Selected Diatomic Flame Radicals*, PURDU-CL-78-06, 1978.
9. Dimpfl, W. L. and Kinsey, J. L.: *Radiative Lifetimes of OH ($A^2\Sigma$) and Einstein Coefficients for the A-X System of OH and OD*, J. Quant. Spec. Radiat. Transfer, **21**, 233 (1979).
10. Dieke, G. AND Crosswhite, H.: *The Ultraviolet Bands of OH: Fundamental Data*, J. Quant. Spec. Radiat. Transfer, **2**, 97, (1963).
11. Huber, K. P., and Herzberg, G.: *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*, Van Nostrand Reinhold Company, 1979.
12. Anderson, W. R., Decker, L. J., and Kotlar, A. J.: *Temperature Profile of a Stoichiometric CH_4/N_2O Flame from Laser Excited Fluorescence Measurements on OH*, Combust. Flame, **48**, 163 (1982).
13. Svehla, R. A. and McBride, B. J.: *Fortran IV Computer Program for Calculation of Thermodynamic and Transport Properties of Complex Chemical Systems*, NASA-TND-7056, 1973.
14. Stufflebeam, J. H. and Eckbreth, A. C.: *CARS Measurements of High Pressure Solid Propellant Combustion*, Combust. Sci. Technol., **66**, 163, (1989).

APPENDIX

Comparison using delta-function or Voigt lineshapes for fitting OH experimental spectra.

In order to verify the validity of using the delta-function approximation for the transition lineshapes, a typical flame spectrum was fitted using both the delta-function and a Voigt lineshape for the transitions. The comparison of the parameters that were determined or fixed in the fit is given in Table 2. The parameters in Table 2 with standard deviations ($\pm 1\sigma$) are those that were allowed to float during the fit. Seven parameters were fitted using the

Table 2 Comparison obtained from fitting a typical flame spectrum using either a Voigt or a delta-function transition lineshape.

NAME	VOIGT LINESHAPE		DELTA-FUNCTION LINESHAPE		DIFFERENCE
	FINAL VALUE	STANDARD DEVIATION	FINAL VALUE	STANDARD DEVIATION	
TEMP	2220.86343	90.74400	2262.57580	92.97000	1.8782%
NTOT	3.91337E+16	1.50E+15	3.73733E+16	1.51E+15	-4.4986%
RCH	550.83341	0.04223	550.82424	0.04242	-0.0017%
RLAM	3089.80000		3089.80000		
INC	0.07514	3.624E-5	0.07514	3.669E-5	0.0009%
SLIT	0.26855	0.00110	0.28312	0.00441	5.4253%
LGTH	0.40000		0.40000		
A0	0.96647	0.00164	0.96616	0.00165	-0.0323%
A1	-1.43822E-5	2.900E-6	-1.45467E-5	2.943E-6	1.1438%
TAU	0.17963	0.01727	N/A		
RERR	1.00000E-4		N/A		
STD	0.00502		0.00508		
XMIN	3077.33927		3077.33941		0.0000%
XMAX	3104.01394		3104.01528		0.0000%
time	~4-8 hrs.		~1 min.		

* The parameter names used in the fitting program are retained in this table: TEMP, temperature (K); NTOT, total number density (molecules/cc); RCH, reference channel corresponding to RLAM, reference wavelength (Å); INC, wavelength increment per channel (Å); SLIT, instrument function FWHM (Å); LGTH, absorption pathlength (cm); A0 and A1, constant and first order baseline coefficients; TAU, Lorentzian FWHM (cm⁻¹); RERR, relative error of the numerical integration; STD, standard deviation of the fit; XMIN and XMAX, minimum and maximum wavelength (Å) of the fitted portion of the spectrum. The approximate time for one iteration of the fit is also given.

delta-function lineshape; an additional parameter, the Lorentzian linewidth of the Voigt profile, TAU, was also fitted when a Voigt lineshape was used for the transitions. Of the seven common parameters fitted in both cases, the two of primary interest are the temperature, TEMP, and the total number density, NTOT. The other five parameters are related to the particular conditions under which the experiment was performed. The largest

differences are associated with those parameters most closely dependent on the value of the transition linewidth, *i.e.* the width of the instrument function (SLIT) and the total number density. The differences in these parameters obtained using the computationally less demanding delta-function lineshape as opposed to the more accurate Voigt lineshape are *ca.* 5% and -4% respectively. A single temperature is obtained from the fit using the Voigt profile since both the rotational temperature, and the translational temperature which determines the Doppler width in the Voigt profile, are assumed to be the same value; also, the pressure broadened (Lorentzian) linewidth is assumed to be the same for all the rotational lines. The approximate time of a least squares iteration for each model is also give in Table 2. Since typically 5-10 iterations are needed to refine a fit, the desirability of using the delta-function model for routine data analysis is apparent.

Navier-Stokes Predictions of Pitch Damping for High L/D
Finned Kinetic Energy Projectiles

Paul Weinacht and Walter B. Sturek
US Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

I. INTRODUCTION

Gun tube launched kinetic energy (KE) projectiles typically fly with some degree of pitching motion caused by launch disturbances such as the whipping motion of the gun tube, inbore balloting of the projectile, and the sabot discard process. The pitching motion of the projectile decreases or damps as the projectile travels downrange due to the aerodynamic properties of the projectile body. Figure 1 shows the pitching motion of an M735 KE projectile as observed in the BRL Transonic Range and is representative of the motion of finned kinetic energy projectiles. This three-dimensional plot shows the horizontal and vertical components of angle of attack, α and β , as a function of the distance downrange. Two-dimensional projections of the angle of attack components as a function of range are also shown. For this projectile, launch disturbances generated initial pitch angles of more than six degrees which are damped as the projectile flies downrange. For finned KE projectiles, the rate at which the pitching motion is damped is a function of the pitch damping aerodynamic coefficient and the body's transverse moment of inertia.

The ability to accurately predict the pitch-damping aerodynamic coefficient of KE projectiles is of significant importance to the projectile designer because the terminal ballistic performance of these projectiles is sensitive to the pitch angle at the target. Small levels of pitching motion may result in significant degradation of the penetrator's terminal ballistic performance. Figure 2 shows measurements¹ of terminal ballistic penetration as a function of pitch angle for a long rod penetrator against a laminated armor target at 65 degrees obliquity. Even at small pitch angles, a significant loss of penetration is observed. If the penetrator and the target are closely matched, this degradation of penetrator performance can result in the inability of the penetrator to defeat the target.

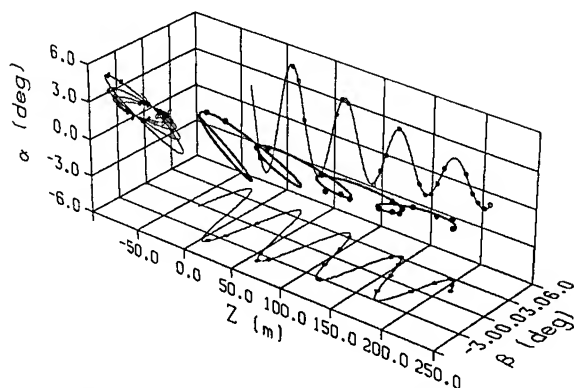


Figure 1. Pitching motion of M735 KE projectile - Round 16423

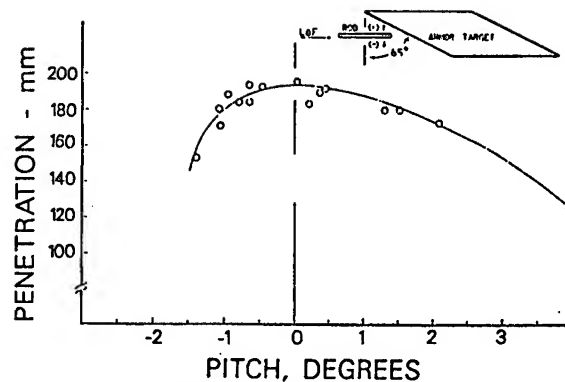


Figure 2. Penetration as a function of pitch angle, 65 degree obliquity target

Fin stabilized KE projectiles typically do not employ active controls for damping the pitching motion, but rely instead on the aerodynamic properties of the body to produce the necessary pitch-damping. A predictive capability for pitch-damping assists in the development of designs which will have acceptable levels of pitch at the target.

By applying linear flight mechanics theory such as that developed by Murphy², it can be shown that aerodynamic side force and moment coefficients acting on a projectile in steady coning motion can be related to the pitch damping force and moment coefficients. Steady coning motion is defined as the motion performed by a missile flying at a constant angle with respect to the free stream velocity vector and undergoing a rotation at a constant angular velocity about a line parallel to the freestream velocity vector and coincident with the projectile center of gravity. This is shown schematically in Figure 3. Coning motion is, in fact, a specific combination of two orthogonal planar pitching motions, plus a spinning motion. It is significant that the combination of these unsteady or time-dependent motions (pitching and spinning) produces a steady coning motion. The use of steady coning motion to determine the pitch damping aerodynamic coefficients provides an interesting and cost effective approach for determining the aerodynamics which are normally associated with unsteady or time-dependent motions.

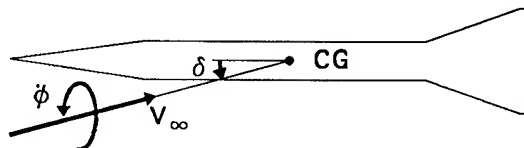


Figure 3. Schematic of coning motion

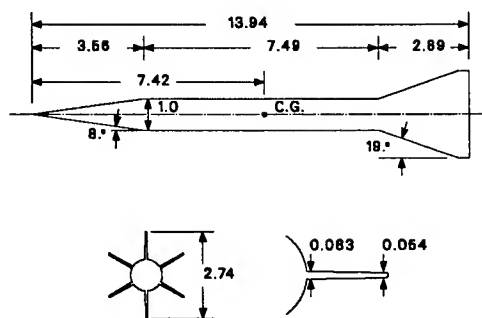
This approach has seen limited application in wind tunnel investigations. Schiff and Tobak³ performed wind tunnel experiments on a conical body undergoing separate or combined spinning and coning motions. Their results showed that, at low angles of attack, the slopes of the side force and moment with angle of attack normalized by the coning rate were in good agreement with predictions of the damping-in-pitch force and moment coefficients obtained using linearized theory. They also demonstrated that the Magnus force and moment (variation of side force and moment with spin rate and angle

of attack) was negligible, thus the linear pitch damping coefficients could be determined from the side force and moment due to coning alone.

As part of the development, Schiff⁴ computed the supersonic inviscid flow about a conical body undergoing coning motion. To compute the flow around the body in coning motion, Schiff made use of a rotating coordinate frame. Within the rotating coordinate frame the flow was steady, thus the steady Euler equations could be solved. The governing equations were modified to include the centrifugal and Coriolis force terms. His computed results compared well with experimental results and with estimates of pitch-damping coefficients using a linear theory. Later studies by Agarwal and Rakich⁵, and Lin⁶ also employed rotating coordinate frames to compute the supersonic viscous flow about conical bodies in coning motion.

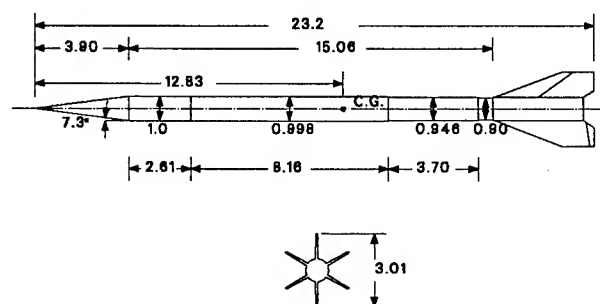
In this report, the supersonic viscous flow field about finned projectiles undergoing coning motion is computed using the Parabolized Navier-Stokes technique of Schiff and Steger⁷. The computations are performed in a rotating coordinate frame similar to that employed in previous investigations. Code modification required to implement the rotating coordinate frame, including the addition of the centrifugal and Coriolis source terms and changes to the shock fitting algorithm are discussed. From the computed flow field, the side force and moment due to coning motion are used to determine the pitch-damping coefficients.

Results are presented in this report for two fielded kinetic energy projectile configurations, the M735 and the M829. Schematics of these projectiles, are shown in Figures 4-5. The fins on both of these projectiles have roll-producing beveled surfaces at the leading and trailing edges of the fins. Particular care has been taken to model the fin geometry accurately. Though not shown, the cylindrical portion of these bodies have a number of circumferential grooves which cover nearly two-thirds of the body. The effect of these grooves is not modeled in the current computations, though it is a subject of current research.



ALL DIMENSIONS IN CALIBERS (ONE CALIBER = 35.2 mm)

Figure 4. Schematic of M735 projectile



ALL DIMENSIONS IN CALIBERS (ONE CALIBER = 27.05 mm)

Figure 5. Schematic of M829 projectile

In this report, the relation between side force and moment coefficients due to coning and the pitch damping force and moment coefficients is outlined. Extensions to the theory required for finned projectiles are described. The computational technique employed to predict the three-dimensional viscous flow field in steady coning motion is then described. Presentation and discussion of the computational predictions and comparisons with range data are made. Finally, the conclusions of this study are presented.

II. THEORETICAL BACKGROUND

In this section, the relation between coning motion and pitching motion is first described. Following this description, the relation between the side force and moment due to coning motion and the pitch damping coefficients is developed.

1. Relation between Coning and Pitching Motions

As was discussed previously, steady coning motion is defined as the motion performed by a missile flying at a constant angle with respect to the free stream velocity vector (angle of attack) and undergoing a rotation at a constant angular velocity about a line parallel to the freestream velocity vector and coincident with the projectile center of gravity. With respect to a non-rolling coordinate frame, the vertical and horizontal components of the angle of attack, α and β , vary in a periodic fashion as the projectile rotates about the free-stream velocity vector. This is shown in Figure 6. The total angle of attack, $\alpha_t = \sqrt{\alpha^2 + \beta^2}$ is constant, however. Both of these components of the angle of attack when plotted as a function of time are seen to be sinusoidal, constant amplitude, pitching motions which are out of phase with each other by one quarter of a cycle, as shown in Figure 7. In addition to the pitching motion, the projectile body will undergo a rotation with respect to the non-rolling coordinate system, which by definition, is a spinning motion. The change in the orientation of a set of axes fixed to the body as the body

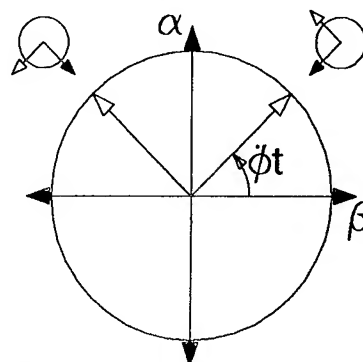


Figure 6. Coning motion with respect to non-rolling coordinates

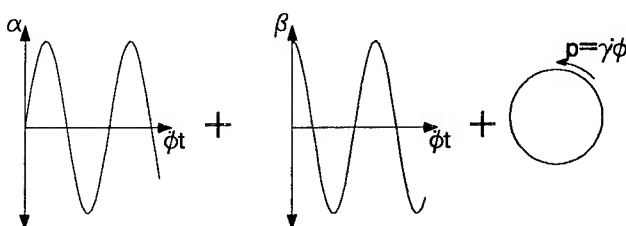


Figure 7. Components of coning motion

performs the coning motion is also shown in Figure 6. The spin rate, p , is proportional to the coning rate, $\dot{\phi}$, as shown below.

$$p = \gamma \dot{\phi} = \cos(\alpha_t) \dot{\phi} \quad (1)$$

Thus, coning motion is a combination of two orthogonal planar pitching motions, plus a spinning motion. These are the three motions shown schematically in Figure 7.

The combination of the two pitching motions and the spinning motion when viewed by an observer fixed to any part of the projectile body produces a motion which does not vary with time. The same can be said for an observer in the flow field external to the body whose position is fixed with respect to the body. Thus, the coning motion and the flow field produced by the coning motion are steady (or non-time varying) phenomenon. This allows steady flow modeling techniques to be applied to determine the flow field produced by this motion. The governing equations must be modified, however, because the coordinate system is rotating. This is discussed in a subsequent section.

2. Relation between Side Moment due to Coning and Pitch Damping Moment

The moment expansion for a finned kinetic energy penetrator in the non-rolling coordinate frame is shown below. This moment expansion is a variant of the expansion discussed by Murphy for symmetric missiles². The difference is that the expression here includes a side moment due to angle of attack, C_{n_α} . The moment formulation uses complex variables to separate the moment components, \tilde{C}_m and \tilde{C}_n which produce rotations in the vertical and horizontal planes.

$$\tilde{C}_m + i\tilde{C}_n = \left[\left(\frac{pl}{V}\right)C_{n_{p\alpha}} + C_{n_\alpha} - iC_{m_\alpha}\right]\tilde{\xi} - \frac{i}{\gamma}[C_{m_q} + \gamma C_{m_{\dot{\alpha}}}] \tilde{\xi}' \quad (2)$$

In the moment formulation, the pitching moment coefficient, C_{m_α} , and pitch damping moment coefficient, $C_{m_q} + \gamma C_{m_{\dot{\alpha}}}$, produce moments proportional to the complex angle of attack, $\tilde{\xi}$, and angular rate, $\tilde{\xi}'$, respectively. The Magnus moment coefficient, $C_{n_{p\alpha}}$, accounts for a side moment due to flow asymmetries from a combination of spin and angle of attack.

The side moment due to angle of attack, C_{n_α} , is retained to account for a side moment variation with angle of attack which is caused by the beveled fins. The existence of this side moment was revealed in a previous study⁸. The mechanism for this moment can be explained by examining the wind and lee-side fins. The fins, which are beveled to produce roll, may experience different flow on the wind and lee-sides of the body. The lee-side fin may be immersed in the wake created by the body. If the flow is sufficiently different on the wind and lee-sides, the bevels on the wind and lee fins will each produce a different lift force. The vector sum of the forces on these fins can

be non-zero, thereby producing a side moment. Of course, at zero angle of attack, the bevels produce only a roll torque; there is no side moment since the vector sum of the lift produced by the fin bevels is zero.

The moment expansion presented above has been developed for rotationally symmetric missiles. Murphy² has shown that the form of the linear force and moment expansion for a symmetric finned missile with three or more fins should have the same form as for a body of revolution. Effects of the orientation of the fins with respect to the pitch-plane are believed to be small for kinetic energy projectiles. This has been verified by computation for several aerodynamic coefficients⁸. Furthermore, these projectiles typically roll at spin rates which are greater than the pitching frequency. Effects of roll orientation are averaged out since the frequency of the effects of roll orientation is much greater than the pitching frequency.

In order to develop the relation between the side moment due to coning motion and the pitch damping moment coefficient, it is convenient to resolve the moment components in non-rolling coordinates into moment components which cause rotations in and out of the plane of the angle of attack. This relation is shown below. Here, C_m is the in-plane moment (the moment which causes rotation of the body in the plane of the angle of attack), and C_n is the side moment (the moment that causes rotations of the body out of the angle of attack plane). Also shown are relations for the complex angle of attack, angular rate, and spin rate. These relations, valid for steady coning motion, have been simplified from the general case of arbitrary motion.

$$C_m + iC_n = ie^{-i\gamma\dot{\phi}t}(\tilde{C}_m + i\tilde{C}_n)$$

$$\tilde{\xi} = \delta e^{i\gamma\dot{\phi}t} ; \quad \tilde{\xi}' \equiv \frac{d\tilde{\xi}}{d(\frac{s}{l})} = i\delta\gamma\frac{\dot{\phi}l}{V}e^{i\gamma\dot{\phi}t} ; \quad p = \gamma\dot{\phi} \quad (3)$$

The resulting expression for the in-plane and side moments due to coning is shown below.

$$C_m + iC_n = C_{m_\alpha}\delta + i\{C_{n_\alpha}\delta + \delta(\frac{\dot{\phi}l}{V})(\gamma C_{n_{p\alpha}} + [C_{m_q} + \gamma C_{m_{\dot{\alpha}}}])\} \quad (4)$$

As expected, the resulting expression for the in-plane and side moments is independent of time. The in-plane moment results only from the pitching moment, C_{m_α} , while the total side moment consists of contributions from the side moment due to angle of attack, C_{n_α} , Magnus moment, $C_{n_{p\alpha}}$, and pitch damping moment, $[C_{m_q} + \gamma C_{m_{\dot{\alpha}}}]$. The side moment which is produced by the coning motion, $\dot{\phi}$, results from the pitch damping and Magnus moments.

Using equation 4, an expression for the side moment due to coning normalized by

the coning rate can be obtained,

$$C_{n\dot{\phi}} \equiv \frac{\partial C_n}{\partial(\frac{\dot{\phi}l}{V})} = \frac{C_n - C_{n\alpha}\delta}{\frac{\dot{\phi}l}{V}} = \delta(\gamma C_{n_{p\alpha}} + [C_{m_q} + \gamma C_{m_{\dot{\alpha}}}] \quad (5)$$

This relation is similar to that presented by Schiff and Tobak³ for bodies of revolution. However, for the case of the finned projectile with side moments due to the beveling, computations at two separate coning rates (which may include zero coning rate) are required to determine the variation of side moment due to coning. For bodies of revolution, a single computation at a non-zero coning rate is sufficient, since the side moment at zero coning rate is zero.

The relationship between the variation of the side moment with coning rate, $C_{n\dot{\phi}}$, the pitch damping coefficient, $[C_{m_q} + \gamma C_{m_{\dot{\alpha}}}]$, and the Magnus moment coefficient, $C_{n_{p\alpha}}$, is shown in Equation 5. At the present time, determination of the Magnus moment coefficient for finned projectiles using three-dimensional Euler or Navier-Stokes approaches has not been performed. This would involve an unsteady time-accurate calculation which would be very expensive. However, the Magnus moment coefficient is typically much smaller than the pitch damping coefficient for many projectiles. In the case of finned projectiles similar to those examined in this study, this has been confirmed by applying simple inviscid theories to estimate the Magnus moment and pitch damping coefficients⁹. Additional confirmation can be found from ballistic range testing of kinetic energy projectiles. The Magnus moment coefficient can be quite difficult to measure due in part to its small magnitude in relation to the other aerodynamic coefficients and due to the low spin rates that these projectiles experience in flight.

For the case where the pitch damping moment coefficient is much larger than the Magnus coefficient, Equation 5 can be rewritten, as shown below. Because this expression is valid in the linear aerodynamics regime (small angles of attack), the cosine of the angle of attack, γ , has been dropped.

$$[C_{m_q} + C_{m_{\dot{\alpha}}}] = \frac{C_{n\dot{\phi}}}{\delta} \quad (6)$$

A similar expression relating side force due to coning to the pitch damping force and Magnus force can be developed using the same approach as discussed above. The Magnus contribution to the force relationship can also be ignored allowing the side force due to coning to be related directly to the pitch damping force.

III. COMPUTATIONAL APPROACH

Computation of the viscous flow field about the finned projectile configurations was accomplished by solving the thin-layer Navier-Stokes equations using a parabolized Navier-Stokes technique. The computations have been performed in a novel rotating coordinate frame which rotates at the coning rate of the projectile. The fluid flow relative to the rotating coordinate frame does not vary with time, allowing the steady (non-time varying) Navier-Stokes equations to be applied. The solution of the steady Navier-Stokes equations can be performed at a reasonable computational cost. In order to implement the rotating coordinate frame, the governing equations have been modified to include the effect of centrifugal and Coriolis forces. The steady thin-layer Navier-Stokes equations are shown below.

$$\frac{\partial \hat{E}}{\partial \xi} + \frac{\partial \hat{F}}{\partial \eta} + \frac{\partial \hat{G}}{\partial \zeta} + \hat{H} = \frac{1}{Re} \frac{\partial \hat{S}}{\partial \zeta} \quad (7)$$

Here, \hat{E} , \hat{F} , and \hat{G} are the inviscid flux vectors, \hat{S} is the viscous flux vector, and \hat{H} is the source term containing the Coriolis and centrifugal force terms which result from the rotating coordinate frame. Each of these matrices are functions of the dependent variables represented by the vector $q(\rho, \rho u, \rho v, \rho w, e)$, where ρ and e are the density and the total energy per unit volume, and u , v , and w , are the velocity components in x , y , and z directions. The inviscid flux terms are shown below.

$$\hat{E} = \frac{1}{J} \begin{bmatrix} \rho U \\ \rho u U + \xi_x p \\ \rho v U \\ \rho w U \\ (e + p)U \end{bmatrix} \quad \hat{F} = \frac{1}{J} \begin{bmatrix} \rho V \\ \rho u V + \eta_x p \\ \rho v V + \eta_y p \\ \rho w V + \eta_z p \\ (e + p)V \end{bmatrix} \quad \hat{G} = \frac{1}{J} \begin{bmatrix} \rho W \\ \rho u W + \zeta_x p \\ \rho v W + \zeta_y p \\ \rho w W + \zeta_z p \\ (e + p)W \end{bmatrix} \quad (8)$$

where

$$U = u\xi_x; \quad V = u\eta_x + v\eta_y + w\eta_z; \quad W = u\zeta_x + v\zeta_y + w\zeta_z \quad (9)$$

The contravariant velocities, U , V , and W , are functions of the velocity components and the metric terms, (i.e. ξ_x, η_y, ζ_z). The metric terms result from the transformation from the body conforming physical coordinate system to the rectangular computational coordinate system. Use of this formulation allows a large variety of projectile shapes to be considered using a single code. The pressure, p , which appears in the inviscid flux terms, can be related to the dependent variables by applying the ideal gas law.

The viscous flux term, \hat{S} , has been modeled using the thin-layer approach which considers only those terms resulting from flow gradients normal to the body. Definition of these terms, which require no modification due to the rotating coordinate frame,

can be found in Reference 7. The turbulent viscosity, which appears in the viscous matrices, was computed using the Baldwin-Lomax turbulence model¹⁰

The source term due to the rotating coordinate system is shown below. This term contains the Coriolis and centrifugal acceleration terms.

$$\hat{H} = \frac{1}{J} \begin{bmatrix} 0 \\ -2\Omega_c \sin \alpha_t \rho v - \rho \Omega_c^2 \sin^2 \alpha_t (x - x_{cg}) + \rho \Omega_c^2 z \sin \alpha_t \cos \alpha_t \\ 2\Omega_c \sin \alpha_t \rho u - 2\Omega_c \cos \alpha_t \rho w - \rho \Omega_c^2 y \sin^2 \alpha_t - \rho \Omega_c^2 y \cos^2 \alpha_t \\ 2\Omega_c \cos \alpha_t \rho v + \rho \Omega_c^2 \sin \alpha_t \cos \alpha_t (x - x_{cg}) - \rho \Omega_c^2 z \cos^2 \alpha_t \\ (-\Omega_c^2 \sin^2 \alpha_t (x - x_{cg}) + \Omega_c^2 z \sin \alpha_t \cos \alpha_t) \rho u - (\Omega_c^2 y \sin^2 \alpha_t + \Omega_c^2 y \cos^2 \alpha_t) \rho v \\ + (\Omega_c^2 \sin \alpha_t \cos \alpha_t (x - x_{cg}) - \Omega_c^2 z \cos^2 \alpha_t) \rho w \end{bmatrix} \quad (10)$$

The thin-layer equations are solved using the Parabolized Navier-Stokes technique of Schiff and Steger⁷. Following the approach of Schiff and Steger, the governing equations, which have been modified here to include the Coriolis and centrifugal force terms, are solved using a conservative, approximately factored, implicit finite-difference numerical algorithm as formulated by Beam and Warming¹¹. The source term has been treated implicitly. Although this term can be included in either the circumferential inversion or in the normal inversion, including this term in the circumferential inversion simplifies slightly the implementation of the shock fitting boundary conditions.

The computations presented here were performed using a shock fitting procedure reported by Rai and Chaussee¹². This procedure solves the five Rankine-Hugoniot jump conditions, two geometric shock-propagation conditions, and one compatibility equation to determine the values of the five dependent variables immediately behind the shock, as well as the position of the shock. By including the implicit part of the source term due to the rotating coordinate frame in the circumferential inversion, the shock fitting procedure of Rai and Chaussee can be used with only minor modifications, provided the free-stream conditions reflect motion produced by the rotating coordinate frame.

The computational results were obtained using a grid which consisted of 60 points between the body and the shock. In the circumferential direction, gridding was performed over the entire body (360 degrees). Over the forebody, 72 circumferential points were used. On the finned portion of the body, 300 points were used in the circumferential direction. The grid over this part of the body was generated using an elliptic grid generation scheme presented by Rai and Chaussee¹³. On the axisymmetric part of the body, about 50 marching (axial) planes were required for each caliber of body length. Axial grid resolution was doubled over the finned portion of the body. The computations were performed on a Cray-2 supercomputer. Solutions required one and a half to two hours of CPU time.

IV. RESULTS

Computations have been performed to determine the aerodynamics of kinetic energy projectiles in steady coning motion. Results have been obtained for two fielded kinetic energy projectiles; the M735 and the M829. The computations have been performed over a range of Mach numbers ($M = 3.0$ to 5.5), coning rates ($\dot{\phi}D/V = 0$ to $.010$), and angles of attack ($\alpha = 1^\circ - 5^\circ$) for free-flight (sea-level) atmospheric conditions. Comparisons are made with data obtained from range firings. The results for each projectile are discussed separately in the following two sections.

1. Results for the M735

The computed variation of the side moment coefficient with coning rate at Mach 4 and two degrees angle of attack is shown in Figure 8. The variation of the side moment coefficient with coning rate is seen to be linear across the range of coning rates examined here. This range of coning rates is representative of the pitching frequencies experienced by this projectile in flight. At Mach 4, the non-dimensional pitching frequency of the projectile is 0.004. The results also show a small non-zero side moment coefficient at zero coning rate. As was discussed previously, this side moment is due to bevels on the fins⁸. The existence of this side moment at zero coning rate requires that computations be performed for at least two coning rates in order to evaluate the variation of the side moment coefficient with coning rate, $C_{n\dot{\phi}}$.

Figure 9 shows the slope of the side moment coefficient with coning rate as a function of δ (the sine of the angle of attack), at Mach 4. The dashed line displayed on this figure is representative of a linear variation of $C_{n\dot{\phi}}$ with the sine of the angle of attack, δ , across the range of angles of attack examined. The computed results show that at small angles of attack, $C_{n\dot{\phi}}$ varies linearly with δ , but departs from a linear variation as the angle of attack increases.

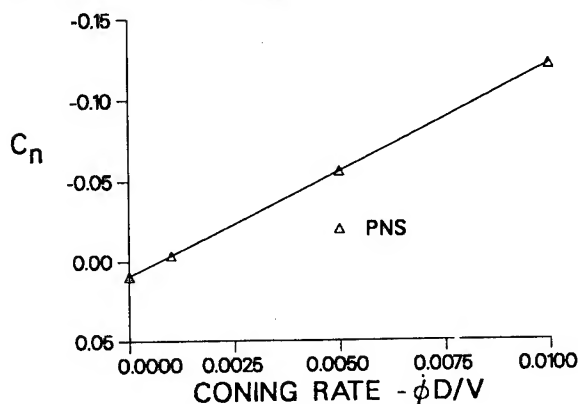


Figure 8. Variation of side moment coefficient with coning rate, M735, Mach 4, $\alpha = 2^\circ$

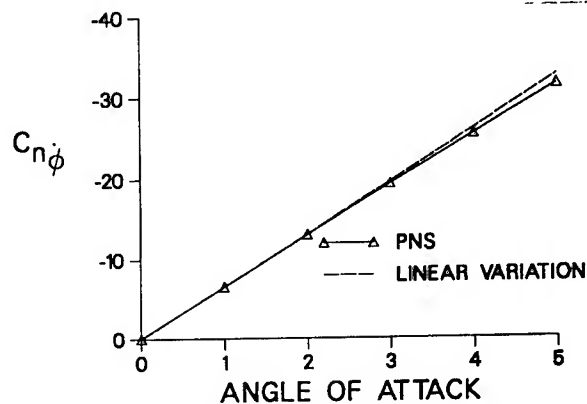


Figure 9. Variation of the slope of the side moment coefficient due to coning, $C_{n\dot{\phi}}$, with the sine of the angle of attack, δ , M735, Mach 4

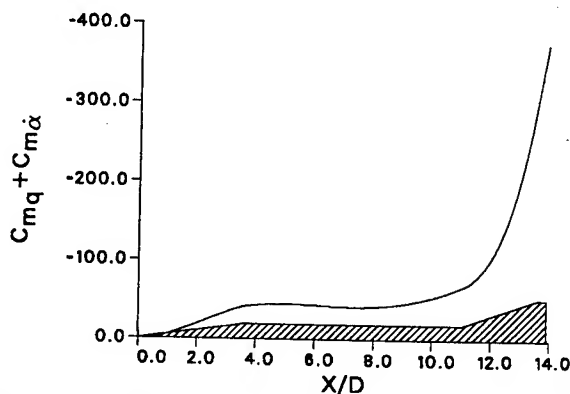


Figure 10. Development of pitch damping moment coefficient over M735 projectile, Mach 4

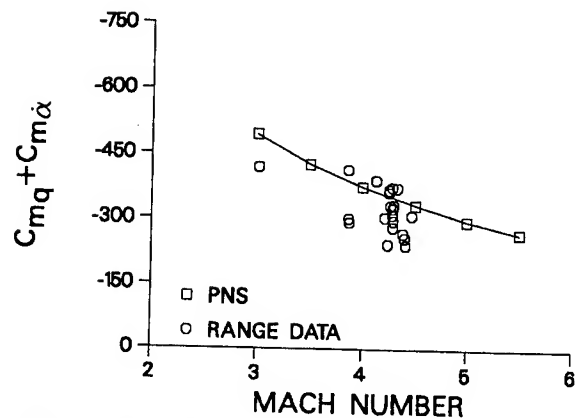


Figure 11. Mach number variation of computed pitch damping moment coefficient compared with range measurements, M735

Figure 10 shows the development of $\frac{C_{n\dot{\phi}}}{\delta}$ over the M735 kinetic energy projectile at Mach 4 and two degrees angle of attack. As was discussed previously, $\frac{C_{n\dot{\phi}}}{\delta}$ should be a reasonable representation of the pitch damping coefficient in the linear aerodynamic regime. This figure shows that the fins contribute most of the side moment due to coning (and hence, the pitch damping) with a smaller contribution from the nose.

The computed variation of $\frac{C_{n\dot{\phi}}}{\delta}$ with Mach number for the M735 is shown in Figure 11. The computed results are compared with range measurements of the pitch damping coefficient. Though the range data shown here is considered well-determined, some scatter is still evident. This is because damping rates are typically difficult to measure. The agreement between computation and range measurement supports the contention that the coning computations allow a reasonable determination of the pitch damping coefficient.

2. Results for the M829

Similar computations were performed for the M829 kinetic energy projectile. Figure 12 shows the computed variation of $\frac{C_{n\dot{\phi}}}{\delta}$ with Mach number for the M829. Again, the computed results are compared with range measurements of the pitch damping coefficients. The range data has considerable scatter because the range firings produced rounds with very low yaw (less than one degree average yaw). Thus, the rate at which the yaw decreased in flight was difficult to determine. The computational results are within the scatter of the range data. Both the computations and the range results show the order of magnitude increase in the coefficient compared with the predictions shown for the M735. This increase is primarily due to the larger length-to-diameter ratio of the M829 ($L/D=23$) compared with the M735 ($L/D=14$).

The predicted variation of the damping force as a function of Mach number is

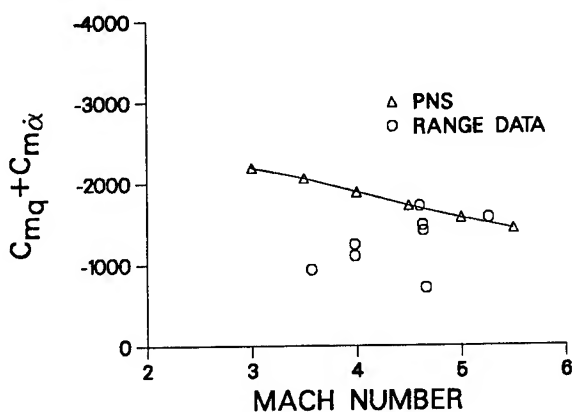


Figure 12. Mach number variation of computed pitch damping moment coefficient compared with range measurements, M829

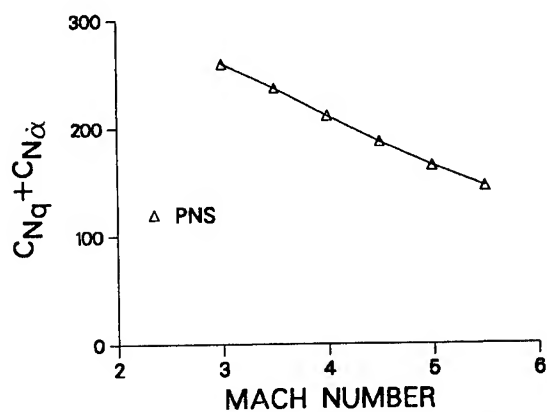


Figure 13. Mach number variation of computed pitch damping force coefficient, M829

shown in Figure 13. This coefficient was determined from the variation of the side force coefficient with coning rate, $\frac{C_{Y_i}}{\delta}$. The pitch damping force coefficient appears in the swerve equation (the equation which describes in-flight motion of the projectile center of gravity). For finned projectiles, the contribution to the swerving motion from this coefficient is very small. Because of this, it is very poorly determined from range firings, thus no experimental data is shown. This coefficient is, however, important for determining the change in the pitch damping moment coefficient due to changes in center of gravity location. Thus, the computed result is significant, particularly since direct measurements from range firings are non-existent.

One situation where changes in the center of gravity location are often encountered is in aerodynamic range testing of kinetic energy projectiles. Because of restrictions on firing kinetic energy projectiles with heavy metal penetrators through aerodynamic test ranges, surrogate projectiles are often used. Externally, these projectiles appear the same as the war round, though internally the heavy metal core has been replaced with a core of a different composition, typically steel. Replacing the penetrator core can result in a shift in the center of gravity. In the case of the M829, this shift is greater than a quarter of a caliber.

In steady coning motion, the projectile rotates about its center of gravity (CG). Flow field computations were performed to determine the effect of CG location on the side force and moment due to coning (and hence on the pitch damping force and moment coefficients). The CG position was moved fore and aft of the baseline CG position by 1 and 2 body diameters. These results are shown in Table 1 for Mach numbers of 3, 4, and 5.

Predictions of the variation of the side force and moment due to coning with center of gravity location can be made using the center of gravity translation relations presented by Murphy². The relations, presented by Murphy for the individual aerody-

dynamic coefficients, are combined to obtain relations for the coefficients of interest here. These are shown below. The coefficients represented with a "hat" represent the predicted value for a CG shift of s_{cg} body diameters. The aerodynamic coefficients on the right side of these expressions represent the values for the baseline configuration.

$$\begin{aligned}\hat{C}_{N_q} + \gamma \hat{C}_{N_{\dot{\alpha}}} + \gamma \hat{C}_{Y_{p\alpha}} &= C_{N_q} + \gamma C_{N_{\dot{\alpha}}} + \gamma C_{Y_{p\alpha}} + s_{cg} C_{N_{\alpha}} \\ \hat{C}_{m_q} + \gamma \hat{C}_{m_{\dot{\alpha}}} + \gamma \hat{C}_{n_{p\alpha}} &= C_{m_q} + \gamma C_{m_{\dot{\alpha}}} + \gamma C_{n_{p\alpha}} - s_{cg}(C_{N_q} + \gamma C_{N_{\dot{\alpha}}} + \gamma C_{Y_{p\alpha}})\end{aligned}\quad (11)$$

By applying equation 5, the above relations are reduced to the following.

$$\begin{aligned}\hat{C}_{Y_{\dot{\phi}}}/\delta &= C_{Y_{\dot{\phi}}}/\delta + s_{cg} C_{N_{\alpha}} \\ \hat{C}_{n_{\dot{\phi}}}/\delta &= C_{n_{\dot{\phi}}}/\delta - s_{cg}(C_{Y_{\dot{\phi}}}/\delta) + s_{cg} C_{m_{\alpha}} - s_{cg}^2 C_{N_{\alpha}}\end{aligned}\quad (12)$$

In deriving these relations, the Magnus force and moment coefficients are retained for consistency, though their effect is thought to be small.

Using these relations and the aerodynamic coefficient predictions for the baseline configuration, predictions of the side force and moment variation due to coning for varying CG position were obtained. These results are also shown in Table 1. The differences between the direct computation of the side moment coefficient at the various CG locations and the values obtained from the CG translation relations is less than 0.2 %, and provides additional validation of the computational approach. The side moment (and hence the pitch damping) varies by more than ± 35 % for a center of gravity shift of ± 2 calibers. Changing the CG position is one possible approach for increasing or decreasing the pitch damping of these projectiles.

Table 1: CG Translation - Comparison of Results Using CG Translation Equations and Direct Computation

MACH NUMBER	s_{cg}	COMPUTED		TRANSFORMED	
		$C_{Y_{\dot{\phi}}}/\delta$	$C_{n_{\dot{\phi}}}/\delta$	$C_{Y_{\dot{\phi}}}/\delta$	$C_{n_{\dot{\phi}}}/\delta$
3.0	-2.0	224.37	-1550.1	224.61	-1551.2
	-1.0	241.89	-1853.2	242.01	-1853.9
	+1.0	276.93	-2564.6	276.81	-2563.7
	+2.0	294.45	-2972.8	294.21	-2970.8
4.0	-2.0	179.59	-1373.4	179.85	-1374.9
	-1.0	195.32	-1615.7	195.45	-1616.6
	+1.0	226.78	-2194.7	226.65	-2193.6
	+2.0	242.52	-2531.5	242.25	-2528.9
5.0	-2.0	192.75	-2080.6	192.46	-2077.8
	-1.0	178.75	-1812.7	178.62	-1811.5
	+1.0	150.78	-1361.1	150.92	-1362.0
	+2.0	136.80	-1177.2	137.08	-1178.8

V. CONCLUSION

The flow field about finned kinetic energy projectiles in steady coning motion has been successfully computed using a parabolized Navier-Stokes computational approach. The computations make use of a rotating coordinate frame. Relative to this coordinate frame, the flow does not vary with time, allowing the steady flow equations to be solved. Using linear flight mechanics theory, the side moment due to coning is related to the pitch damping and Magnus moment coefficients. For small Magnus moment coefficients, the pitch damping coefficient is directly proportional to the side moment due to coning. The computational results show that the side moment due to coning varies linearly with coning rate over a range of coning rates that encompasses the pitching frequency of the projectile. For constant coning rate, the side moment coefficient also shows a linear behavior with the sine of the angle of attack up to about two degrees angle of attack, and deviates slightly from the linear behavior at higher angles of attack. The computational predictions of the slope of the side moment coefficient with coning rate normalized by the sine of the angles of attack have been compared with pitch damping coefficients determined from range firings. For the M735 kinetic energy projectile, the predictions are in good agreement with the range data. The computational results for the M829 are within the scatter of the range data. Both the computational predictions and the range data for the M829 show a substantial increase in the pitch damping coefficient over the M735 results. This is primarily due to the larger length-to-diameter ratio of the M829. The variation of the pitch damping force and moment coefficients with center of gravity location determined from direct computation and by applying center of gravity translation relations. The excellent agreement between the two approaches provides additional validation of the computational approach.

The favorable comparisons with range data and the efficiency of the computational approach demonstrate the utility of the this newly developed capability. To date, several advanced kinetic energy projectile concepts have been examined using this capability.

References

1. Roecker, E., and Grabarek, C., "The Effect of Yaw and Pitch on Long Rod Penetration into Rolled Homogeneous Armor at Various Obliquities," The Proceedings of the 9th International Symposium on Ballistics, Shrivenham, England, May 1986, Volume 2, pp. 467-473.
2. Murphy, C.H., "Free Flight Motion of Symmetric Missiles," U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, Report No. 1216, July

1963. (AD A442757)

3. Schiff, L.B., and Tobak, M., "Results from a New Wind-Tunnel Apparatus for Studying Coning and Spinning Motions of Bodies of Revolution", AIAA Journal, Vol. 8, No. 11, November 1970, pp. 1953-1957.
4. Schiff, L.B., "Nonlinear Aerodynamics of Bodies in Coning Motion", AIAA Journal, Vol. 10, No. 11, November 1972, pp. 1517-1522.
5. Agarwal, R., and Rakich, J.V., "Computation of Supersonic Laminar Viscous Flow Past a Pointed Cone at Angle of Attack in Spinning and Coning Motion", AIAA Paper 78-1211, AIAA 11th Fluid and Plasma Dynamics Conference, Seattle, WA, July 1978.
6. Lin, T.C., "A Numerical Study of the Aerodynamics of a Reentry Vehicle in Steady Coning Motion", AIAA Paper 78-1358, AIAA Atmospheric Flight Mechanics Conference, Palo Alto, CA, August 1978.
7. Schiff, L.B., and Steger, J.L., "Numerical Simulation of Steady Supersonic Viscous Flow," AIAA Journal, Vol. 18, No. 12, December 1980, pp. 1421-1430.
8. Weinacht, P., "Computation of the Aerodynamic Forces and Moments Acting on a Kinetic Energy Projectile at Angle of Attack", U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, Technical Report to be published.
9. Devan, L., "Nonaxisymmetric Body, Supersonic, Inviscid Dynamic Derivative Prediction," AIAA Paper 89-2195-CP, AIAA 7th Applied Aerodynamics Conference, Seattle, WA, August 1989.
10. Baldwin, B.S., and Lomax, H., "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA Paper 78-257, 16th Aerospace Sciences Meeting, January, 1978.
11. Beam, R., and Warming, R.F., "An Implicit Factored Scheme for the Compressible Navier-Stokes Equations," AIAA Journal, Vol. 16, No. 4, 1978, pp. 85-129.
12. Rai, M.M., and Chaussee, D.S., "New Implicit Boundary Procedure: Theory and Applications," AIAA Paper 83-0123, Reno, Nevada, January 1983.
13. Rai, M.M., and Chaussee, D.S., "Calculation of Viscous Supersonic Flows over Finned Bodies," AIAA Paper 83-1667, Danvers, MA, July 1983.

Setting Enlistment Standards on the ABLE to Reduce Attrition (U)

Dr. Leonard A. White*, Mr. Roy D. Nord, & Dr. Fred A. Mael
U.S. Army Research Institute
5001 Eisenhower Avenue, Alexandria, VA 22333-5600

There is mounting evidence that enlisted job performance is multidimensional. Success on the job is the joint product of skills and knowledges (i.e., "can-do" components) and "will-do" components of effort, discipline, and adaptability. The Services use the Armed Services Vocational Aptitude Battery (ASVAB) and educational attainment as primary preenlistment tools for managing performance after enlistment. The ASVAB strongly predicts skilled performance, while being a high school diploma graduate (HSDG) does not predict this aspect of soldier performance as well. In contrast, being a HSDG is strongly related to discipline and attrition during the first term of enlistment, while ASVAB is less strongly related to these performance components.

The Army has developed a new test known as the Assessment of Background and Life Experiences (ABLE) that can add to the ASVAB and HSDG as predictors of job performance. The ABLE is a temperament scale developed through an extensive, theoretically based process. A review of the research on temperament scales and biographical data revealed a variety of measures that showed validity for predicting important aspects of job performance. In a parallel effort, subject matter experts estimated the likely relationship between these temperament constructs and 72 different criteria of successful soldiering in 19 representative Military Occupational Specialties (MOS).

Based on the results of the literature review and expert ratings, preliminary temperament measures were developed. An initial battery containing the most promising variables was administered to 11,000 trainees in four MOS. Guided by the results from the Preliminary Battery, 11 new measures of temperament were developed to form the ABLE. Four validity scales were developed as a check on the quality of information from these 11 content scales. Presently, the 11 content scales are arrayed to form three primary predictor composites: Achievement, Discipline, and Stress Tolerance.

In 1985, the resulting form of ABLE was administered during the Army's Project A Concurrent Validation (CV) research to 9,430 first term enlisted personnel at 15 CONUS posts and in USAREUR. A Longitudinal Validation (LV), in which nearly 50,000 receptees were administered ABLE in late 1986 and early 1987, is still in progress. Results from the CV research show that the ABLE provides incremental validity over the ASVAB and HSDG as a predictor of performance on the job. The Achievement composite has the

highest incremental validity as a predictor of leadership potential and job effort. The Discipline composite is most predictive of indiscipline: First term soldiers scoring in the lowest third on ABLE-Discipline had 300% more Articles 15 and Flag Actions than those scoring in the upper third. Although we also expected ABLE to predict attrition, the cross-sectional CV data did not allow evaluation of that hypothesis. In this paper, we examine the incremental validity of the ABLE as a predictor of first term attrition using data from the LV sample.

Factors Influencing Army Attrition

Attrition of first term recruits continues to be a concern for the Army. It adds substantially to the cost of recruiting and training, limits the pool of candidates for NCO leadership roles, and has a disruptive effect on unit cohesion and morale.¹ Researchers have investigated a number of factors that influence attrition during the first term of enlistment. These include personal characteristics, such as educational credentials, gender, age, race, and organizational variables, such as one's enlistment term and time in the Delayed Entry Program (DEP).² The most important of these variables is educational attainment. The first term attrition rate for high school graduates is about 24%, and for non-high school graduates about 50%. Attrition rates vary by MOS, but are typically higher for females and for recruits with low scores on the Armed Forces Qualification Test (AFQT). Blacks and other non-whites have been found to have lower attrition rates than whites, though differences are not always statistically significant.³

Information regarding predictors of attrition is used by the Army to reduce total attrition by modifying recruiting missions and job enlistment standards. In recent years, the Army has attempted to reduce attrition by limiting the recruitment of non-high school graduates and by taking more recruits above the 50th percentile on the AFQT. Still, roughly a third of recruits fail to complete their term of enlistment.⁴

Missing from previous models of attrition are measures of individual differences in temperament. Many trainee discharges are attributed to poor adjustment, lack of effort, or lack of discipline. Based on a review of the temperament literature⁵, we hypothesized that recruits who are stress resistant and disciplined, as measured by the ABLE, would better adapt to the rigors of military life, and hence be more likely to complete their term of enlistment.

Plan For This Report

The focus of this paper is to develop an approach for setting enlistment standards on the ABLE to reduce first term attrition. First, we describe the LV database used in this research. Then, we use logistic regression analysis to estimate the degree to which ABLE improves prediction of first term attrition over HSDG, ASVAB, and other factors. Next, we evaluate alternative strategies for using the ABLE in selection by comparing the savings from reduced attrition with the increased recruiting costs that result from ABLE screening. In the final section, we discuss potential uses of the ABLE in manpower allocation to reduce training costs and improve performance.

Estimating Attrition Effects

Database. The sample consisted of soldiers in 21 MOS who completed a battery of new cognitive and non-cognitive tests (including ABLE) during the LV phase of the Army's Project A. The ABLE testing took place on the first day of in-processing at the reception station. Attrition information for cases in the sample was obtained from automated enlistment records and merged with the LV database. The records of soldiers who separated from the enlisted force either to become officers or because of reasons beyond their own control (e.g. accidental death) were excluded from the analyses. The total remaining sample consisted of 42,733 soldiers.

Independent Variables in the Attrition Model. Table 1 presents the mean and standard deviation of the explanatory variables used in the estimation. All of these variables are categorical except ABLE, AFQT, and months in the DEP. Some further explanation of the ABLE scales and the AFQT score is provided below. In the total sample, 10% were female; 72% were white; 11% had two year enlistment tours; 51% were assigned to a combat MOS; and, 90% were HSDGs. The expected direction of the effect of the explanatory variables on attrition is shown in Table 1.

ABLE. The ABLE form used in the LV research contained 199 items that form 11 content scales and four response validity scales. The content scales form three primary composites called Achievement, Discipline, and Stress Tolerance. Achievement-oriented persons work hard, set high standards of personal achievement, and seek positions of leadership and influence over others. Disciplined persons are dependable, accepting of authority, and obey rules and regulations. Those high in Stress Tolerance display emotional stability and cope well under stress.

The Department of Defense has incorporated the ABLE into a biodata and temperament instrument called the Adaptability Screening Profile (ASP). The ASP has two scored sections: a section with 50 items from the Armed Services Applicant Profile (ASAP), and a section with a 70-item version of the ABLE. The abridged form of the ABLE developed for the ASP measures the primary composites of the original ABLE with high reliability and minimal loss of validity⁶.

The ASP was developed for the primary purpose of reducing attrition during soldiers' first term of enlistment. The research presented in this paper was motivated by the requirement to recommend enlistment standards to reduce attrition for the ABLE portion of the ASP. Therefore, although examinees responded to all 199 items, only the 70 items on Form A of the ASP were scored. ABLE Form A composites for Achievement, Discipline, and Stress Tolerance were computed. Then, these composite scores were added with unit weights to form an overall ABLE score, which is reported as a percentile with a mean of 50. The ABLE percentile score has high internal consistency reliability ($\alpha=.92$), and is highly correlated ($r = .97$) with its counterpart on the 199-item form.

The ABLE Validity scale was developed to detect inaccuracy in examinees' responses caused by intentional or subconscious attempts to look good. The mean and standard deviation of scores on this scale indicate only minimal dissembling by examinees in responding to the ABLE. Importantly, the research presented in this paper provides a first time

Table 1
Independent Variables in the Attrition Model

Independent Variable	Mean	SD	Description	Expected Effect
Female	.10	--	female=1; male=0	+
White	.72	--	white=1; nonwhite=0	+
Two Yr	.11	--	2 yr term=1; 3-4 yr=0	-
DEP	3.77	3.17	Months in DEP	-
HSDG	.90	--	HSDG=1; nongrad=0	-
GED	.03	--	GED=1; nongrad=0	-
Combat	.51	--	combat=1; noncmbt=0;	+
AFQT	56.13	19.93	AFQT percentile	-
Valid	17.02	3.40	ABLE Validity scale	?
ABLE	49.42	28.90	ABLE percentile	-

opportunity to determine if the Validity scale can improve the predictive accuracy of the ABLE composite score against the criterion of attrition.

Armed Forces Qualification Test (AFQT). The ASVAB is comprised of 10 subtests and is used to predict the trainability and job proficiency of applicants for the Armed Services. The AFQT score is a composite of four ASVAB subtests. It is reported as a percentile score with a mean of 50.

Relation of ABLE to First Term Attrition

Logistic regression analysis was used to relate attrition to the independent variables. Attrition, the dependent variable, was coded as a dichotomous variable with a 1 defined as separated before 21 months of service and 0 defined as still active after 21 months.

The binary logistic regression model is appropriate for this analysis because it assumes that attrition is a discrete event the probability of which varies between 0 and 1. The independent variables in the model are used to estimate the probability of attrition for an individual given that individual's characteristics. This relationship can be expressed as:

$$P(\mathbf{x}_i) = \text{Prob}(\text{Attrit}_i=1) = 1/(1+e^{-\beta\mathbf{x}_i}), \quad (1)$$

where \mathbf{x}_i is a vector of individual characteristics, $P(\mathbf{x}_i)$ is the probability that the i th individual will separate during the first 21 months, and the β_j 's are the parameters to be estimated relating the independent variables to the likelihood of attrition. Thus, the dependent variable (attrition) is modeled as a function of several variables

simultaneously. In this way, the effect of variation in ABLE scores on attrition rates is isolated.

Table 2 shows the correlations among the explanatory variables in the model. The ABLE composite score had uniformly low correlations with the other independent variables in the model, indicating that the ABLE provides new information to the Army for managing attrition and performance. Notably, the ABLE is uncorrelated with HSDG. The unexpected negative correlation between the AFQT and the ABLE Validity scale suggests greater dissembling by examinees in the lower Mental Categories.

Table 2
Correlation Among the Independent Variables

Var	Fem	Wht	2yr	DEP	HSD	GED	Cmb	AFQT	Valid	ABLE
Female	1.00									
White	-.07	1.00								
Twoyr	.02	.09	1.00							
DEP	.00*	.17	.09	1.00						
HSDG	.11	-.09	.12	.14	1.00					
GED	-.06	.03	-.06	-.08	-.52	1.00				
Combat	-.34	.04	-.02	-.07	-.07	.02	1.00			
AFQT	-.03	.39	.26	.19	-.06	.04	-.03	1.00		
Valid	-.02	-.13	-.09	-.01	-.01	-.01	.03	-.19	1.00	
ABLE	.02	-.10	-.04	-.08	.00*	-.02	.04	.09	.34	1.00

* $p > .05$ (n.s.). For all other coefficients, $p < .001$

Two logistic regression models were examined to evaluate the incremental contribution of ABLE to the prediction of attrition. Table 3 presents the results of the tests of these two models. All of the explanatory variables were statistically significant at the 0.01 level. The difference in twice the logarithm of the likelihood ratio for the two models is distributed as a chi-square that can be used to test the incremental validity of the ABLE over prediction provided by the remaining explanatory variables. This chi-square test is highly significant, $X^2(4)=498.22, p<.0001$, indicating that ABLE adds significantly to the prediction of attrition in the model. The degrees of freedom for this chi-square are equal to the difference between the number of free coefficients being compared.

The probability of attrition increased as a nonlinear function of lower ABLE scores. The predicted attrition rate for ABLE scores at the 1st, 10th, and 20th percentiles was, 34%, 26%, and 21%, respectively. For ABLE scores above the 30th percentile the attrition rate stabilized between 16% and 17%. The higher attrition rates among those scoring at the lowest ABLE percentiles compares with the higher attrition among non-high school graduates. In the sample, the attrition rate for non-high school graduates exceeded the rate among HSDGs by 15 percentage points.

In line with prior research, females were found to have higher attrition rates than males by 12 percentage points; whites (as compared to non-whites) had a higher probability of attrition by 7 percentage points; and, attrition was 7 percentage points higher for recruits who enlisted for 3 or 4 years. In addition, the probability of attrition was somewhat higher for those recruits who spent a shorter time in the DEP, had lower AFQT scores, were assigned to combat MOS, and had higher scores on the ABLE Validity scale.

Table 3
Logistic Regression Coefficients for Two Models
With Attrition as the Dependent Variable

Independent Variables	MODEL 1	MODEL 2
Intercept	-0.4018 (.0625)	0.5929 (.1050)
Female	0.7689* (.0417)	0.8083* (.0421)
White	0.5320* (.0315)	0.4512* (.0315)
Two Year Enlistment	-0.4248* (.0475)	-0.4675* (.0479)
DEP	-0.1226* (.0163)	-0.1265* (.0164)
DEP ²	0.0061* (.0014)	0.0061* (.0014)
HSDG	-0.9158* (.0420)	-0.9034* (.0423)
GED	-0.1892* (.0721)	-0.2194* (.0729)
Combat MOS	0.2239* (.0273)	0.2467* (.0276)
AFQT	-0.0078* (.0007)	-0.0055* (.0007)
Valid	--	0.0188* (.0041)
ABLEPCTL	--	-0.0517* (.0040)
ABLEPCTL ²	--	0.0008* (9.62x10 ⁻⁵)
ABLEPCTL ³	--	-3.9*10 ⁻⁶ * (6.5x10 ⁻⁷)
N	42733	
Mean Attrition Rate	19.9	
Model Chi-Square	1655.12* (9df)	2153.34* (13df)

* p < .01.

Note. Standard errors are shown in parentheses.

Cost Effective Selection Standards

The purpose of the following discussion is to provide a brief overview of the methodology used to first determine, and then compare, the costs and benefits of alternative ABLE standards. The basic concept employed to determine the optimal ABLE selection standard is simple: reject applicants for whom the expected cost of attrition is larger than the cost of screening an additional applicant. For the purposes of this analysis, we focused exclusively on the use of ABLE as a screen to reduce first-term attrition; additional benefits of the ABLE screen, such as reduced disciplinary problems and improved job performance, were ignored, primarily because of the difficulty in quantifying these benefits in terms that could be compared to the costs of screening additional applicants. As a result, our recommendations with respect to minimum selection standards are conservative -- that is, if the ancillary benefits of ABLE screening were considered, higher minimum standards than those recommended here would be justified.

The difficulties in implementing this procedure have to do with obtaining reasonable estimates of expected attrition and replacement costs, and developing a workable algorithm for evaluating the difference between them under alternative selection policies. We make a number of simplifying assumptions to do this. These will be noted as they become relevant in the following discussion.

Attrition Probabilities. The specifics of the basic attrition model have already been described. We begin here with a more general model than the one actually estimated to clarify the assumptions inherent in that model and show how the actual estimates were used in the cost-benefit analysis.

Denote the probability that an individual (i) with a vector of observable characteristics \mathbf{x}_i will be an attrition in job (j) as p_{ij} . Then

$$p_{ij} = f(\mathbf{x}_i, \beta_j) \quad (2)$$

represents a set of job specific equations relating applicant characteristics to the probability of attrition in each job. The first simplification we make is to eliminate the subscript j in this equation. While previous research suggests that there are important differences across jobs in both mean rates and in the effect of certain individual characteristics on attrition, the data to estimate separate equations for all Army jobs using ABLE is not currently available. Furthermore, for purposes of initial selection, this simplification has no effect, as long as the assignment of applicants to specific jobs is not affected by changes in the minimum selection standard.

This simplification leads to equation (1), denoted here as

$$p_i = f(\mathbf{x}_i, \beta), \quad (3)$$

where f is the logistic regression function, \mathbf{x}_i is the vector of explanatory variables in the regression, and β is the set of coefficient estimates.

In order to use the predicted probabilities to evaluate alternative minimum standards, two additional steps are needed. While it would be theoretically possible to calculate a different p_i for each applicant and make eligibility dependent directly on the predicted attrition probability, this approach would be impractical within the Army's current selection and assignment system. One alternative to the direct use of predicted probabilities is to assume that all elements of x , with the exception of the ABLE score, are at their means, and derive a single minimum ABLE score for all applicants. This approach is undesirable not only because of large and systematic differences in the predicted probabilities among recruits who differ on other elements of x (High School attainment, for instance), but also because there are significant differences in the replacement costs among groups of applicants. The approach we take is a compromise between these two extremes. We carry out a separate analysis for each of seven major categories of recruits, where the criteria defining the categories are those used by the U.S. Army Recruiting Command to define major segments of the recruiting market ("mission boxes") -- high school attainment, AFQT category, and gender.

The effect of this treatment of individual differences on equation (2) is that we partition x_k into three components indexed by mission box $k=\{1,2, \dots 7\}$, the first, denoted a , contains the ABLE score, the second, y_m includes the variables that define mission box (m) and the third, z_m , contains the means for mission box m of the other variables in the attrition equation. This yields an equation that can be used to predict the mean attrition probability for each possible ABLE score for each mission box:

$$p_{mk}=f(a_k, y_m, z_m) \quad (4)$$

where $k=\{1,2,\dots 99\}$, indexes ABLE percentile scores.

The final step in obtaining the attrition rate prediction needed for the analysis is to generate mean predicted rates for accepted and rejected applicants for each potential ABLE cutoff score. This is simply

$$p'_{km} = \int^k f(a, y_m, z_m) da \text{ for rejectees, and} \quad (5)$$

$$p^*_{km} = \int^{\infty} f(a, y_m, z_m) da \text{ for acceptees.} \quad (6)$$

In practice, we approximated these integrals by simply summing the mean predicted rates within percentile for each potential cut score. To reduce unnecessary computation outside the range of feasible cutoff scores, we limited our analysis to rejection rates of 0-60%.

Attrition Costs. Attrition imposes significant costs on the Army, including some that are not easily quantified. These include disruptive effects of personnel turbulence and the discipline-related problems that often precede attrition. These are not accounted for in our analysis, which focuses only on two directly measurable costs associated with first-term attrition -- excess training costs, and increased recruiting needs.

For this analysis, we treat all training costs for a recruit who separates within the first 21 months of service (the time frame used for the attrition model) as wasted resources. Most of the attrition over this period occurs either during or shortly after completion of training. Even

when attrition occurs later, the portion of the training investment that is recouped through subsequent productive service is relatively small. For attrition that occurs early in training, only a small portion of total training cost is saved, since most of these costs are not responsive to changes in class size after a class has started.

Attrition increases recruiting requirements, because in order to maintain fixed levels of trained and deployed manpower, soldiers who fail to complete their tour must be replaced, resulting in increased recruiting missions.

Currently, expected attrition rates are factored into both recruiting missions and training requirements. Thus, a reduction in expected attrition rates due to improved selection is likely to take some time to be reflected in reduced accessions and training loads. In this sense, our analysis is a "steady state" approach, reflecting the savings that accrue after the system has adjusted to the lower loss rates.

While it is not necessary for this analysis to model MOS-specific differences in attrition probabilities, it is essential to account for the wide variation in training costs across MOS. This is because (a) training costs vary widely across MOS (ranging from \$8800 to over \$46,000, with a mean of roughly \$16,200 in our sample); and (b) the distribution of mission boxes within MOS also varies significantly. This means that average training costs differ from one mission box to another. Therefore, in calculating the expected training costs for each mission box, we use a weighted average of MOS-level training costs, where the weights are the 1986 proportions of the mission box allocated to each MOS. The resulting estimates of mean training costs range from a low of \$15,100 (for GFB's) to a high of \$17,500 (for GFA's). The data used to compute training costs for each MOS was extracted from the ATRM-159 report compiled by the Resource Management Office of the U.S. Army Training and Doctrine Command (TRADOC).

Replacement Costs. The costs of using a selection device include not only the minimal costs of actually administering the selection instrument, but also the cost of attracting and processing the applicants required to replace those screened out by the device. While in some cases these costs may be negligible, for the Army they are not. A substantial proportion of the recruiting budget is spent to attract applicants, and, for the most desirable recruit categories, these costs escalate at an increasing rate as the total number of needed applicants rises. Thus, any policy that increases the required number of these applicants will increase not only total recruiting costs, but also the average cost per applicant. In addition to this effect, estimation of replacement costs must account for the substantial processing costs incurred prior to the determination of an applicant's eligibility under a proposed selection device. In the case of ABLE, this determination will probably be made only after the recruit reaches the Military Enlistment Processing Station (MEPS). At this point, most of the costs of a signed contract will already have been incurred. It is not feasible to determine a precise value for the proportion of total recruiting costs that should be charged to a rejected applicant. For the purposes of this analysis, we assume a proportion of 90%.

In order to estimate either replacement costs or the value of reductions in recruiting loads due to reduced attrition, reasonable

estimates of recruiting costs are required. For these, we have relied on estimates produced by Kearn and Nord⁷ for their analysis of cost-effective recruit selection policy. Their approach varies from previous work in this area primarily in that they develop estimated non-constant marginal cost curves for all categories of recruits. Previous research has generally assumed that costs are constant for all groups other than high-school graduates in the upper half of the AFQT distribution (mission boxes GMA and GFA). Recent evidence from the difficult recruiting year of FY89 suggests that this assumption may not be valid. To obtain baseline costs for this analysis, we used the Kearn and Nord estimates to calculate marginal costs for each mission box for the 1990 projected recruiting missions. The resulting marginal costs are as follows: HSDG, female, AFQT percentile 50-99 (GFA), \$19,200; HSDG, female, AFQT 31-49 (GFB), \$7,422; High school graduate, male, AFQT percentile 50-99 (GMA), \$18,800; HSDG, male, AFQT percentile 31-49 (GMB), \$8,807; HSDG, male, AFQT percentile 16-30 (GM4), \$3,800; GED, male, AFQT percentile 50-99 (HMA), \$7,200; and non-high school graduate, male, AFQT percentile 50-99 (NMA), \$7,200. Replacement costs for each potential ABLE cut score were calculated by determining the net change in the number of applicants needed under that cut score, calculating the new marginal recruiting cost, and multiplying the product by the adjustment factor of .90. Marginal costs for high quality applicants were assumed to increase at the rate of 2% for each 1% increase in the number of applicants required. Costs for all non-quality groups except GM4 were assumed to increase at the rate of .5% for each 1% increase in applicants.

Determining the Optimal Cut Score. Given the data on predicted attrition, and recruiting and training costs, the procedure for obtaining the optimal selection standard is straightforward:

(1) Calculate p_{km}^* , and p_{km}' for each ABLE percentile score, for each mission box.

(2) Calculate the rejection rate (r_{km}) for each ABLE screen. (Note: The rejection rate for a given percentile cut score varies slightly among mission boxes, because of differences in the distributions of ABLE scores.)

(3) Calculate the number of avoided attritions for each mission box at each cut score as $a_{km}^* = (p_{km}' - p_{km}^*)r_{km}M_m$, where M_m is the accession mission for mission box m.

(4) Calculate the attrition adjusted recruiting mission $M_m^* = M_m - a_{km}$

(5) Calculate the net replacement cost $c_{km} = .9g(M_m^*)r_{km}M_m^*$, where $g(M_m^*)$ is the marginal recruiting cost for mission box m at a mission of M_m^* .

(6) Calculate net savings $S_{km} = t_m a_{km} - c_{km}$, where t_m is average training costs for mission box m.

(7) The optimal cut point is determined by selecting the k that yields the smallest non-negative S_{km} .

Results

Table 4 shows expected attrition rates among accepted and rejected

Table 4
Expected Attrition Rates Under Different ABLE Cut Scores

ABLE CUT (%tl)	GFA	GFB	GMA	GMB	GM4	HMA	NMA
0	24.3 (n.a.)	29.6 (n.a.)	14.3 (n.a.)	18.8 (n.a.)	21.2 (n.a.)	30.6 (n.a.)	37.2 (n.a.)
1	24.2 (44.5)	29.5 (46.9)	14.3 (26.7)	18.8 (38.6)	21.1 (40.3)	30.4 (48.8)	37.0 (61.0)
3	23.8 (41.0)	29.2 (47.3)	14.1 (26.9)	18.5 (33.1)	21.0 (36.1)	29.7 (48.0)	36.5 (55.7)
5	23.5 (39.6)	29.1 (46.8)	13.9 (26.4)	18.2 (32.2)	20.4 (34.9)	29.5 (46.5)	36.0 (55.2)
10	22.8 (38.0)	28.5 (42.6)	13.4 (24.3)	17.5 (30.3)	19.6 (33.1)	29.0 (44.9)	35.3 (52.2)
15	22.4 (35.9)	27.8 (39.6)	13.1 (22.3)	16.9 (28.1)	19.2 (31.5)	28.6 (42.0)	34.6 (48.6)
20	22.2 (33.2)	27.5 (38.7)	12.9 (21.0)	16.4 (26.7)	19.0 (30.1)	28.2 (39.5)	34.3 (47.0)
25	21.9 (31.7)	27.3 (37.5)	12.7 (19.8)	16.1 (25.1)	18.8 (28.7)	27.8 (38.0)	34.1 (44.9)
30	21.8 (30.8)	27.2 (36.0)	12.6 (18.8)	15.8 (23.9)	18.8 (27.4)	27.7 (35.8)	34.1 (43.8)
35	21.7 (29.7)	27.2 (36.0)	12.5 (18.1)	15.7 (23.1)	18.6 (26.3)	27.6 (34.7)	34.0 (42.4)
40	21.7 (28.1)	27.2 (33.8)	12.4 (17.4)	15.6 (22.3)	18.5 (24.8)	27.7 (33.8)	34.0 (41.3)
45	21.7 (27.4)	27.2 (32.4)	12.4 (16.8)	15.6 21.7	18.5 (24.0)	27.7 (33.4)	34.0 (41.1)

Note. Predicted rates for rejectees are shown in parentheses.

Key: G:High school graduate; H:GED; N:Non-graduate
F:Female; M:Male
A:AFQT I-III A; B:AFQT III B; 4:AFQT IV

Table 5
Optimal ABLE Cut Scores by Mission Box:
Attrition and Recruiting Effects

MISSION BOX	CUT SCORE	ATTRITION RATE OF ACCEPTEES	ATTRITION RATE OF REJECTEES	NUMBER REJECTED	PERCENT REJECTED	1990 CONTRACT MISSION	ESTIMATED NET SAVINGS (\$000'S)
GFA	0	24.3	n.a.	n.a.	n.a.	9200	n.a.
GFB	5	29.1	46.8	135	3.4	4010	97
GMA	0	14.3	n.a.	n.a.	n.a.	51000	n.a.
GMB	0	18.8	n.a.	n.a.	n.a.	22891	n.a.
GM4	34	18.6	26.5	516	32.0	1616	682
HMA	8	29.1	45.1	158	7.7	2059	131
NMA	16	34.5	48.3	325	15.6	2084	453

recruits (p'_{km} and p^*_{km}) by mission box for a set of alternative ABLE cut scores. To clarify the information presented in Table 4, consider the effect of ABLE screening on attrition among GFAs. Based on our analysis, a 3% ABLE screen is expected to reduce attrition for this mission from 24.3% to 23.8%. The expected attrition rate among applicants that would be rejected by this screen is 41.0%.

As can be seen in Table 4, ABLE screening can reduce attrition in every mission. For ABLE cutoff scores at or below the 30th percentile, the expected attrition rates are typically 50%-75% higher for rejectees, as compared to acceptees. This percentage is slightly less than the difference in the attrition rate of GMAs as compared to non-high school graduates.

Optimal cuts for each mission box are shown in Table 5, along with our estimates of the net savings produced by each cut. The expected percentage of otherwise qualified applicants that would be rejected by the proposed screens is shown in column 6 of Table 5. In every case, the percentage of rejectees is slightly less than the value of the optimal ABLE cut score. This is due to the near zero correlation of HSDG, ASVAB, and gender, with the ABLE. For the GFB, HMA, and NMA missions, the expected attrition rate among the rejectees exceeds 45%. In contrast, the high recommended cutoff for GM4s is not due to especially high attrition, but rather to low recruiting costs. For GFAs and GMAs, the marginal cost of recruiting is more than \$18,000, which exceeds the variable cost of training. For these missions, the attrition rates of rejectees would need to be greater than 50% before recommending cuts, because the cost of attrition is less than that of replacements.

The total annual savings from using the ABLE for selection is estimated to be \$1.36M. If this savings is less than the additional costs of the total Army share of ABLE testing, our analysis would suggest that ABLE is cost-effective, provided that the minimum recommended cutoffs are enforced.

Use of ABLE in Classification

The analyses presented above focus on the use of ABLE in selection. One additional option is to use the ABLE in classification, to account for MOS differences in training costs and possible differential prediction of attrition by the ABLE at the MOS level. The advantage of a classification is that savings can be achieved without incurring significant increases in recruiting costs.

One classification procedure for using the ABLE to reduce training costs would be to allocate ABLE-identified high attrition risks to MOS with lower training costs. In addition, there is some indication that the relationship of ABLE composites to attrition varies by job. If this differential predictability holds, it suggests a strategy for using the ABLE in selective placement. Specifically, overall system performance could be improved by allocating recruits who are lower in achievement orientation and stress resistance to jobs where these attributes are less related (or possibly unrelated) to performance; those with higher scores would be needed in occupations where stress tolerance and discipline are stronger factors in overall job performance.

REFERENCES

1. Hosek, J. R., Antel, J., & Peterson, C. (1989). Who stays, who leaves? Attrition among first-term enlistees. Armed Forces & Society, 15, 389-409.
2. Buddin, R. (1988). Trends in attrition of high-quality military recruits. The RAND Corporation, R-3539-FMP.
3. Baldwin, R. H., & Daula, T. V. (1984). The cost of high-quality recruits. Armed Forces & Society, 11, 96-114.
4. Hosek et al., *ibid*.
5. Kamp, J. D., & Hough, L. M. (1988). Utility of temperament for Predicting job performance. In L.M. Hough (Ed.), Literature Review: Utility of temperament, biodata, and interest assessment for predicting job performance. ARI Research Note 88-02, Alexandria VA.
6. White, L. A. (1989). Development of Forms A and B of the Assessment of Background and Life Experiences (ABLE). Presentation to the Joint Services Selection and Classification Working Group. San Diego, CA.
7. Kearl, C. E., & Nord R. D. (1990). How Much Soldier Quality? Cost Effective Recruit Selection Policy for a Smaller Army. Paper to be presented at the Army Science Conference, Durham, NC.

Predicting the Decay of Mobile Subscriber Equipment (MSE)
Operator Skills (U)

Dr. Robert A. Wisher and Dr. Mark A. Sabol*
U.S. Army Research Institute
for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Unless practiced, job skills deteriorate after initial training. This fact has long been established in the military training literature^{1,2} and is the justification for unit sustainment training. The extent of skill decay depends on many factors, among them individual differences in aptitude, degree of initial learning, and duration of the no-practice period, or retention interval. This paper examines the contribution that a broader, more conceptual understanding of a task might have on skill decay.

Skill retention, the inverse of decay, denotes the remembrance of how a skill should be performed long after it has been initially acquired. In the purest sense, skill retention refers to remembrance without intervening practice. The existing model for predicting skill retention is one which the Army Research Institute (ARI) developed³ and TRADOC published as the User's Manual for Predicting Military Task Retention⁴. Use of this model has been demonstrated to predict accurately the decay of simple procedural skills⁵. Examples of such skills are boresighting a weapon, donning a gas mask, or disassembling a rifle. The present research describes the initial stages of a program to develop a model that can predict the decay of more complex skills. These are the skills that support the operation and maintenance of many newly fielded Army systems, such as the Mobile Subscriber Equipment communication system.

Of the factors that influence how well a skill is remembered, the method of initial skill training is critical⁶. During initial training, the learner begins to compose a representation of knowledge based on both classroom instruction

and on prior knowledge. For example, classroom instruction might detail the sequence of steps necessary to initialize a piece of electronic communication equipment. A student's prior knowledge of electrical systems, however, can help maintain memory for this sequence. Knowing that power-up must occur before a key-in step can be executed is a simple example of how prior knowledge applies a constraint to learning. The structure of the knowledge acquired during initial training assumes a form that can be later reinforced, networked, modified, or perhaps forgotten.

Training can, in part, be regarded as simply the addition of new information to memory. The correct and timely recall of this information in the context of a task is identified as competent performance. The way this knowledge is structured, however, has a strong influence on its accessibility⁷. In a popular view of the development of knowledge structure⁸, production system models of learning regard skill development as a step-wise progression from a declarative form to a compiled, proceduralized form that is characteristic of skilled performance.

The transition of knowledge from the declarative to compiled representation is, we believe, a key to the retention of a skill over prolonged periods of non-use. One indicator that this transition has been accomplished is a student's internal reference to a comprehensive conceptual understanding of the dynamics of a system's operation. A conceptual understanding, also referred to as a mental model, is invented by the student or conveyed through a learning strategy as an appropriate representation of the target system. Appropriateness is measured in terms of accuracy, consistency, and completeness. A good mental model is easily retained and serves as an effective aid in reconstructing details and applying constraints to a task at hand.

Since conceptual understanding might be unique to the knowledge and learning style of each particular student, it is not possible to propose that a single representation applies to all students. This can lead to difficulties in measuring adequately the depth and nature of any individual's conceptual understanding. Still, we consider it useful to explore the prospects of relating conceptual understanding to skill retention, as the payoff to Army training could be large. The remainder of this paper describes the initial research that addressed this issue.

This research was conducted on the skills required to operate the Mobile Subscriber Equipment, or MSE, system. Based on the current retention model, predictions were made on the retainability of most MSE operator skills. An MSE Knowledge Test was developed and administered to active and reserve component users. This written test attempted to measure the depth of understanding the concepts underlying the operation of MSE. After a retention interval of several months, performance tests on the operation of selected MSE procedures were conducted in the field. Relationships between our measure of conceptual understanding and retention performance were then derived. These steps are described below.

Mobile Subscriber Equipment

MSE is a "high-tech" radio telephone system designed to provide secure communications over a five-division corps area. MSE is complex technology and its optimal use requires the smooth performance of complex procedural skills by its operators. In addition, MSE was acquired as a non-development item (NDI), in order to allow its rapid fielding, and will eventually be used by both the Active Army and the Reserve Component. These factors have given rise to a number of personnel selection and training concerns. Chief among these concerns is the extent to which critical operator skills will be retained over extended periods without practice.

The MSE acquisition is, therefore, an opportunity for conducting timely and useful skill retention research. The results of this research will serve two purposes: 1) to predict retention of MSE operator skills, in particular, and 2) to expand an existing retention model to cover more complex skills, in general.

Phase I: Prediction of Retention

The application of the current model to MSE skills, which are considered to be complex and appropriately cognitive in nature, is the first phase in evaluating the model's applicability to such skills. The next step will be to determine the actual retention of these skills over various periods and to compare actual performance with that predicted.

The existing model was applied as instructed in the User's Manual to most MSE operator procedures, producing a predicted retention curve for each of 85 skills. Of these, 33

are skills to be performed by MSE Network Transmission Operators (MOS 31D) and 52 are to be performed by MSE Network Switch Operators (MOS 31F).

Six subject matter experts (SMEs) were identified within a signal battalion at Fort Gordon, Georgia, by recommendation of their company commander. Three of these SMEs were expert on the MOS 31D procedures and three were expert on MOS 31F procedures.

Candidate procedures, generally referred to as tasks, were obtained from two sources: 1) inventories generated during the development of the Soldier's Manual by the Area Communication Department, Ft. Gordon, and 2) lists of critical skills generated during field evaluation of MSE by the U. S. Army Operational Test and Evaluation Agency (OTEA)⁹. From these, a total of 85 procedures were chosen to be rated: 33 from MOS 31D and 52 from MOS 31F. These procedures include all the "operate" tasks from the Signal Center inventories (but only a few of the "install" and none of the "maintain" tasks). Likewise, all of OTEA's operator skills were represented in the final list of rated procedures.

The intention of the knowledge test was to measure an operator's conceptual understanding of the MSE system. A preliminary version of a test of MSE knowledge was compiled by sampling from existing tests and by generation of new items. In addition, items were generated with the express purpose of tapping a soldier's understanding of the workings of a total MSE system (that is, the way in which the various assemblages of a working MSE network function together to maintain communication). This initial version was submitted to instructors at Fort Gordon for technical review and was then pilot tested on small groups of Fort Gordon soldiers.

Pilot testing eliminated those items which failed to discriminate, by producing either no errors or all errors among the subjects. From the remaining pool, items were drawn so that the range from low to high difficulty was covered evenly. Superimposed on this sample, items with high difficulty were deliberately added. This was done to promote maximum variability among soldiers' scores on the new test, since the intention was to determine the extent to which variance in test scores predicts variance in retention.

The resulting MSE Knowledge Test consisted of 8 sections:

- Acronyms: Give the meaning for each of 18 MSE related acronyms (e.g., "RAU = Radio Access Unit");
- Fill-ins: Provide 26 missing terms or numbers in sentences describing MSE ("Each RAU has _____ (how many?) RT-1539s?");
- Matching I For each of 10 pieces of MSE equipment, choose the best description of its function ("RAU - bridges MSRT and Node Center.");
- Functions For each of 20 stated MSE functions, name of the piece of equipment which provides it ("Radio coverage of local terrain - RAU.");
- Matching II Match 10 MSE terms with their definitions ("Full duplex - channel for simultaneous 2-way communication);
- Deducible Directory Given the three necessary tables, deduce the directory number for each of 5 subscribers ("Node Commander, 2^d Platoon, A Company, 3rd Division Signal Battalion.");
- Network Show (by filling in terms and sketching connections) how a given set of units, antennas, and cables could be formed into a working MSE network (27 possible points).
- Components Given a diagram representing the connection between a subscriber and a switch, match 8 components in the chain with its name ("Switch Entry Panel").

The MSE Knowledge Test had a maximum score of 133 points.

All SMEs indicated that they were confident of their expertise. A brief explanation of the instruction was then given, beginning with a quick reading of the ten questions to be asked about each task. These ten questions concerned 1) the existence and 2) the quality of any job or memory aid available to the soldier during performance of the procedure; 3) the number of steps in the procedure; 4) whether the steps in the procedure must be performed in a strict order; 5) whether

built-in feedback is provided as each step is performed correctly; 6) whether a "GO" rating depends on completing the performance of the procedure within a time limit; 7) how complex the mental processing requirements of the procedure are; 8) how many facts must be kept in memory as the procedure is being performed; 9) how hard it is to memorize those facts; and 10) how much fine motor control the procedure requires.

The MSE Knowledge Test was administered in various-sized groups in several settings, including a small classroom, a large lecture hall, and a motorpool garage. The instructions noted that there were questions on the test which all soldiers were not expected to be able to answer, since the topics covered were outside their particular MOS; however, all were encouraged to do their best all on questions. No time limit was put on completion of the test, but all finished within 90 minutes.

Results

For each of the ten questions in the manual's interview protocol, the three responses obtained from the three different SMEs were converted to one response by determining the median. The ten median scores were then summed to obtain a total score for each procedure; this sum constitutes the first data point for a given procedure. A second data point was obtained for each procedure by calculating the sum over the last eight interview questions, that is, excluding those two pertaining to the existence and quality of a job or memory aid. For all of the MSE procedures listed, this job or memory aid turned out to be the soldier's Technical Manual (TM), which is easily accessible inside the shelter which houses the MSE equipment. There were, therefore, two scores obtained for each procedure--one "with TM" and one "without TM."

Using the prediction tables in the User's Manual, these total scores were then converted to the percentages of soldiers predicted to receive a "Go" after various retention intervals. These results were then charted¹⁰. Each chart shows predicted retention over a twelve-month period in two curves, one for the situation in which the soldier uses the TM and one in which the procedure is performed without it. Figure 1 shows two examples.

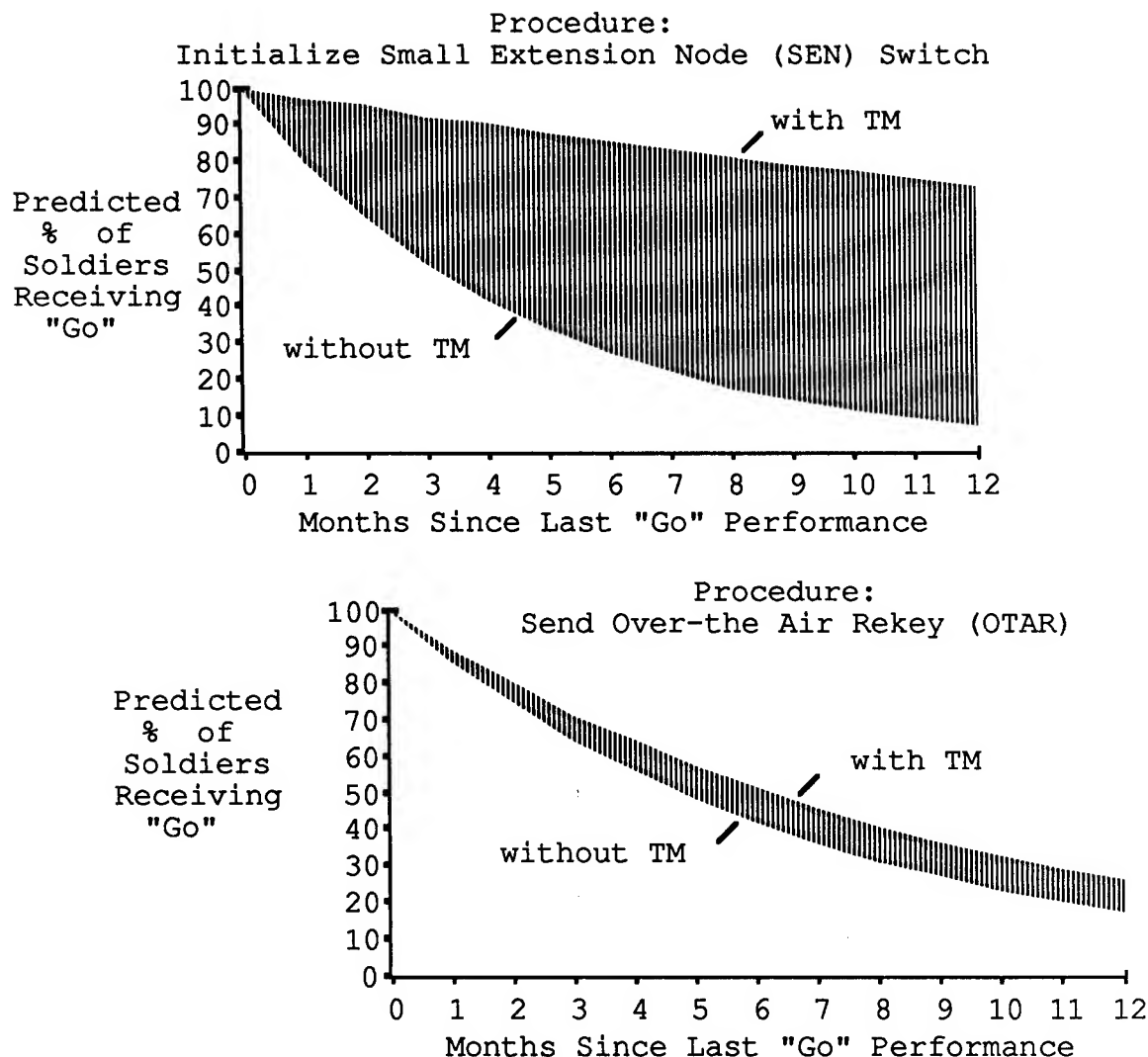


Figure 1. Predicted retention curves for two MSE procedures.

From these charts, the procedures were then categorized as follows: "well-retained"--those procedures for which greater than 75% of soldiers were predicted to receive a "Go" two months after training (assuming that the soldiers can use their TMs); "moderately-retained"--those procedures for which 50-75% of soldiers (using their TMs) were predicted to receive a "Go" after two months without practice; and "poorly-retained"--those for which less than 50% of soldiers using TMs were predicted to

receive a "Go" at 2 months. Figure 2 indicates that a majority of the MOS 31D procedures were predicted to be well retained, while MOS 31F was predicted to have a more even distribution of procedures retained poorly, moderately, and well.

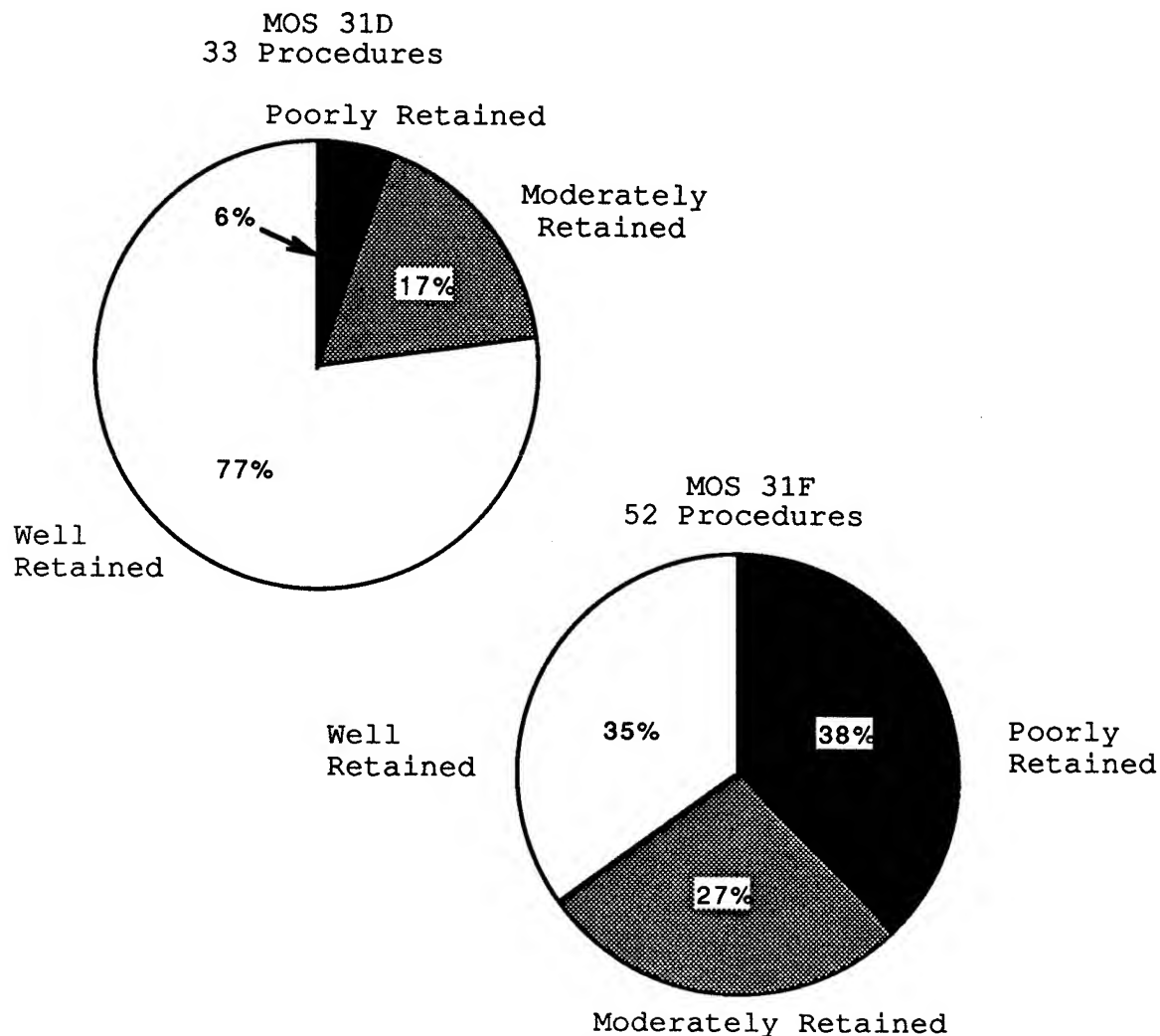
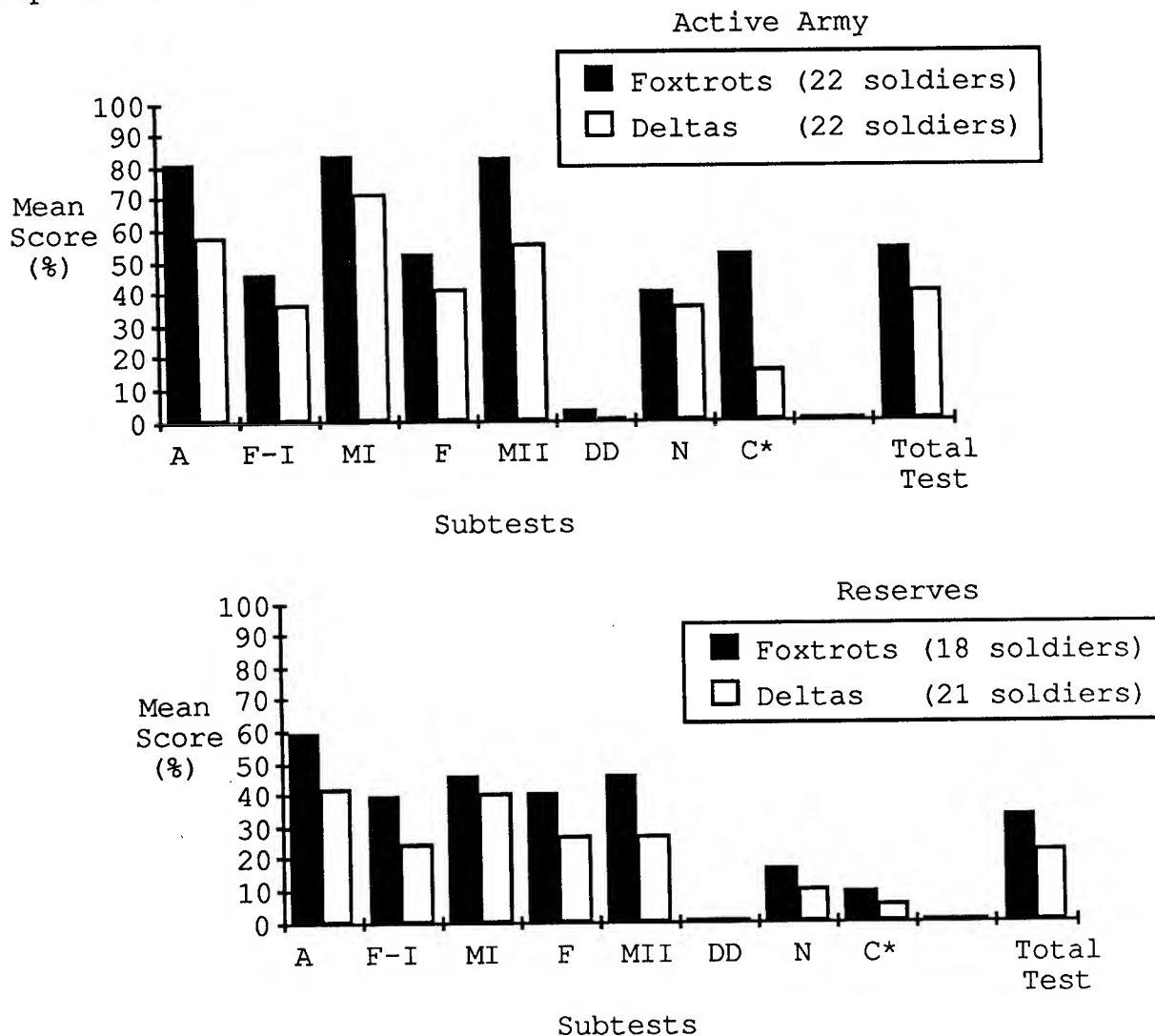


Figure 2. Percentages of the procedures rated which were predicted to be easy to retain, moderately retainable, and difficult to retain, separately for each MOS.

As intended, the knowledge test proved difficult. Within the active unit, tested upon completion of formal training, the

maximum score was 73%; within the reserve unit, tested ninety days after formal training, the maximum score was 65%. Figure 3 depicts the results.



*Subtest names:

A=Acronyms F-I=Fill-ins MI=MatchingI F=Functions MII=MatchingII DD=Deducible Directory N=Network C=Components

Figure 3. MSE Knowledge Test results from Fort Gordon and from Arkansas National Guard.

Note that within both samples of soldiers, active and reserve, the Switch Operators (31F) scored higher than the Transmission Operators (31D) on all subtests. This superiority apparently reflects the involvement of switch operators in more network interactions, resulting in better understanding of MSE network functioning. We consider this highly consistent result to be support for the general validity of our test.

Phase II: Retention Measures

The next phase of the research assessed the validity of these predictions and sought to determine the relationship between conceptual understanding and skill retention. This research is continuing, so the preliminary results are reported here. In this phase, soldiers' proficiency on a sample of the MSE operate procedures are being measured after several months with little or no practice have elapsed since initial training. Proficiency immediately after training is assumed to be at 100% (i.e., 100% of the soldiers receive a "GO" rating), since the "GO" rating defines successive completion of training for each soldier. The change in proficiency between 100% and the retention measurement will provide an empirical value for the retention of each procedure. Such values will then be used to assess the validity of the current model as a predictor of the retainability of MSE operator skills.

Retention measures have been obtained from small samples of soldiers on four MSE procedures. Although this sample of procedures is not yet sufficient to allow evaluation of the validity of the model's predictions, even this small sample showed interesting effects of conceptual knowledge upon retention. Therefore, the details of the methodology are reported below.

A total of sixty soldiers have been tested on performance thus far. Each subject was previously tested on the MSE Knowledge Test. The determination of "GO" or "NO GO" on a procedure were provided by senior NCO's familiar with the operation of the MSE.

Each soldier was individually directed to perform a task, such as initialize a small extension node, with or without the TM, as they preferred. ARI researchers recorded performance as described by the rater. Included in this data were times to complete substeps of the task, or procedure, and the overall "GO" or "NO GO" performance determined by the rater.

Results

One of the four procedures investigated thus far, "Send over-the-air rekey" was retained well enough that all eight 31F soldiers tested received a "GO" rating. The other three procedures all produced roughly equal numbers of "GOs" and "NO GOs." We could therefore ask whether a soldier's conceptual understanding (as measured by our MSE Knowledge Test) had an influence on the likelihood of that soldier's retaining the trained skills needed for those three procedures. Figure 4 shows the MSE Knowledge Test scores obtained by those soldiers receiving "GO" and "NO GO" ratings. For all three procedures, higher test scores were associated with "GO" ratings.

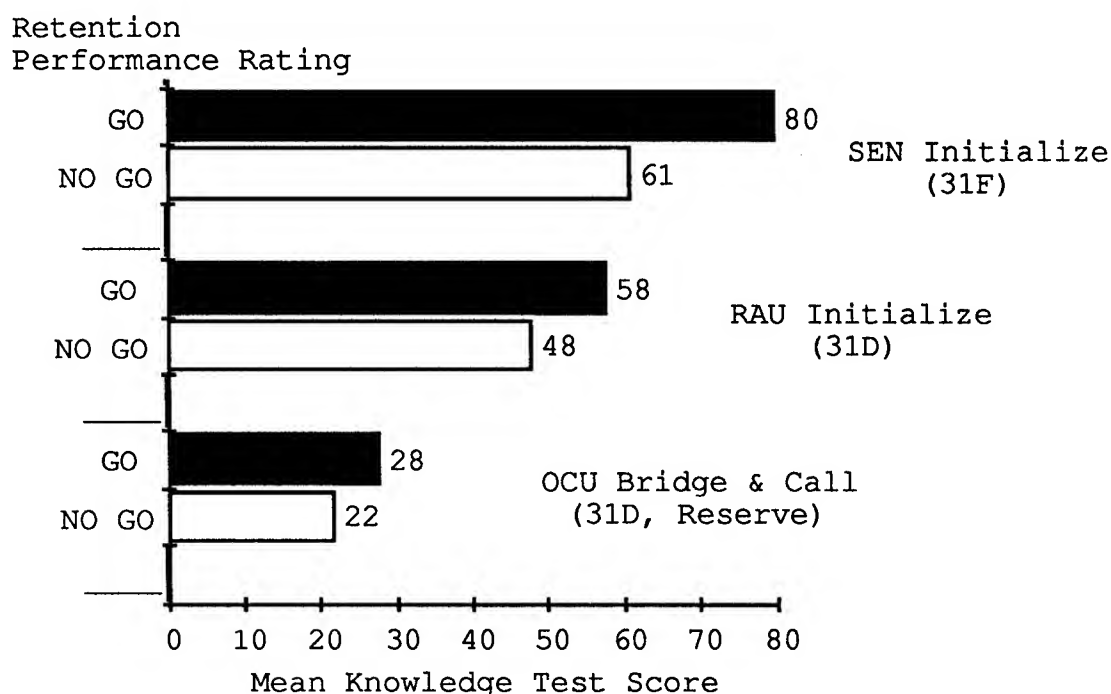


Figure 4. Relationship between conceptual knowledge and retention of three MSE procedures.

Three one-tailed t -tests were performed on the data, comparing the mean test scores for "GO" soldiers to the mean test score for "NO GO" soldiers, separately for each procedure.

For one procedure, SEN Initialization, the effect just reached significance at the .05 level ($t(7)=1.90$); for another, OCU Bridge and Call, the effect was marginally significant ($t(7)=1.48$, $.05 < p < .10$). For the third procedure, RAU Initialization, the effect just failed to reach significance at the .10 level ($t(12)=1.34$, $p < .12$). We find it encouraging that all three statistical tests approached significance despite the small samples involved.

Discussion

The preliminary results lend support to our hypothesis that conceptual understanding of a system aids in remembering the details of its operation. The extensive data collection planned later for newly trained MSE operators will enable us to answer this question more convincingly. With the addition of time as a dependent measure, we will be better able to determine accessibility of remembered information, and isolate points in a procedure where soldiers begin to have difficulties.

An important contribution of this approach might be in the determination of individual soldiers who can benefit most from sustainment training. The current model of skill retention predicts overall performance in a unit rather than by individual. If the model predicts, for example, that 50% of a unit will receive a "No Go" after three months of no practice, it would be helpful to know which half of the unit might need refresher training, in light of constraints on the training budget.

As with other skills, MSE operator skills decay if not practiced. The relationships between conceptual understanding and skill retention which we are beginning to understand can help to reduce this decay through alterations in the classroom training routines. By better understanding the deeper concepts underlying the components of a communication system, the functions of the various extension nodes, the flow of a signal through an assemblage, and the networking arrangements possible, a soldier can more effectively interrelate the details of operating such a system. The effectiveness of automatic switching within a tactical communication system depends upon the skills of its operators.

ENDNOTES

1. Vineberg, R. (1975). A study of the retention of skills and knowledge acquired in basic training. ARI Technical Report 75-10.
2. Shields, J. L., Goldberg, S. L., and Dressel, J. D. (1979). Retention of basic soldiering skills. ARI Research Report 1225.
3. Rose, A. M., Hagman, J. D., Radtke, P. H., & Shettel, H. H. (1985). User's manual for predicting military task retention. ARI Research Product 85-13.
4. TRADOC Form 321-R.
5. Rose, A. M., Czarnolewski, M. Y., Gragg, F. E., Austin, S. H., Ford, P., Doyle, J., & Hagman, J. D. (1985). Acquisition and retention of soldiering skills. ARI Technical Report 671.
6. Farr, M.J. (1987). The long-term retention of knowledge and skills. New York: Springer-Verlag.
7. Gentner, D. and Stevens, A. (Eds) (1983). Mental models. Hillsdale, NJ: Lawrence Erlbaum Associates.
8. Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.
9. Garretson, J. F. (1988). Independent evaluation report (interim) of the Mobile Subscriber Equipment (MSE) follow-on operational test and evaluation (FOT&E) August-October 1988. OTEA Report IER-FO-1148.
10. Sabol, M., Chappell, L., and Meiers, C. (1990). Predicted decay of Mobile Subscriber Equipment (MSE) operator skills. ARI Research Product, 90-11.

Oxidative Decontamination of the Chemical Nerve Agent, VX (U)

Yu-Chu Yang, Dr.(1), Linda L. Szafraniec, Mrs.(2),
William T. Beaudry, Mr.*(2), and Dennis K. Rohrbaugh, Mr.(2)
(1) Physical Protection Directorate, (2) Research Directorate
U.S. Army Chemical Research, Development and Engineering Center
Aberdeen Proving Ground, Maryland 21010-5423

I. INTRODUCTION.

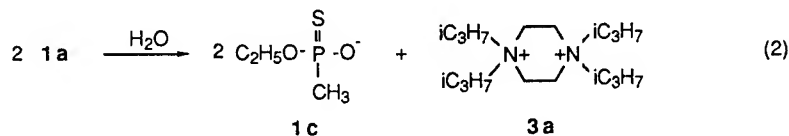
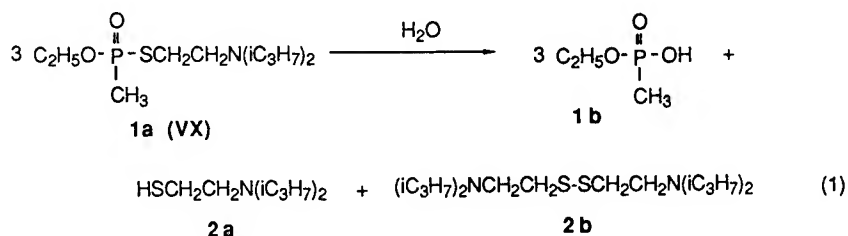
VX, S-(2-diisopropylaminoethyl) O-ethyl methylphosphonothiolate, 1a, is a highly toxic chemical nerve agent first synthesized in the mid-1950's and currently stockpiled by the U.S. Army. Although VX has been known for almost 40 years, its chemical nature is far more complicated than the phosphonofluoridate ester analogs, and only limited information on the detoxification of bulk quantities of VX exists. The standard decontaminant, DS2, is used to detoxify VX under combat conditions.¹ While extremely effective at destroying the agent, DS2 has deleterious effects on some materials.² In addition, since DS2 is considered to be a hazardous material, the resulting solutions are classified as hazardous waste and must be regulated in accordance with the statutes of RCRA (the Resource Conservation and Recovery Act). Consequently, the use of DS2 to detoxify R&D quantities of VX as well as its use in the large-scale demilitarization of leaking and/or obsolete agent-filled munitions is undesirable since it would generate large quantities of regulated hazardous waste.

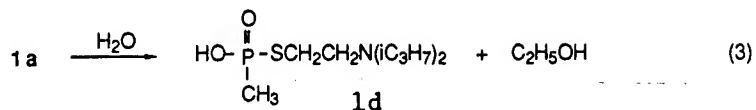
During the last decade, CRDEC has become increasingly aware of its responsibilities regarding the safe, effective decontamination of its chemical agents and the lawful storage and disposition of the resulting decontaminated waste. In Jan 1986, the State of Maryland enacted regulations that identified certain chemical warfare (CW) agents and their decontaminated residues as hazardous waste. CRDEC responded by preparing a package documenting the protocols used to detoxify the agents as well as the chemical analyses and the toxicity data for the decontaminated residues.³ The result was that on 24 Feb 1989, the Attorney General of the State of Maryland approved these decontamination procedures and certified that the decontaminated residues were not listed as hazardous waste from specific sources.⁴

Using high field NMR and advanced GC/MS techniques, research studies are currently underway to develop new methodologies for deliberate decontamination of CW agents by soldiers in the field. These investigations, which involve studies into the fundamental chemistry of chemical agent decontamination, are pointing the way to better, safer and more cost-effective methods for detoxifying CW compounds. This paper describes one such study.

II. BACKGROUND.

Most of the toxic organophosphorus esters can be detoxified quickly by hydrolysis in alkaline solutions.^{5,6} However, relative to the G-agents (Sarin and Soman), the hydrolysis of the phosphonothiolate esters [RP(O)(OR')(SR'')] is much slower even at very high pH values.^{7,8} The estimated half-life of the spontaneous hydrolysis of VX, 1a, (³¹P: δ 61.7, water) was 80 hrs at 20°C in pure water in the absence of any pH control.⁹ In addition, multiple hydrolysis pathways have been reported.^{8,9} As shown in Eqs. 1-3, VX hydrolyzes via simultaneous cleavage of the P-S, S-C and P-O bonds to form a series of products. Although both the ethyl methylphosphonic acid, 1b (³¹P: δ 30.2, water), and the O-ethyl methylphosphonothioic acid, 1c (³¹P: δ 76.4, water), are relatively non-toxic, the S-(2-diisopropylamino)ethyl methylphosphonothioic acid, 1d (³¹P: δ 43.4, water), is almost as toxic as VX (LD50, iv rabbit, mg/kg: VX=0.008; 1d=0.017).¹⁰ Contrary to the findings by Epstein et al. that VX hydrolyzed with the single reaction path shown in Eq. 1 at pH values above 10,⁸ 22% 1d and 78% 1b were detected by ³¹P NMR from the reaction of 0.05M VX with 2N NaOH in an aqueous solution of 10 vol% 2-propanol (necessary to solubilize VX under alkaline conditions). This almost precludes hydrolysis as an effective reaction to detoxify VX.





Currently, the Standard Operating Procedure (SOP) approved by the State of Maryland for the conversion of VX into a non-hazardous waste specifies the use of an ethanolic HTH (high test hypochlorite) mixture with emphasis on the minimization of the ethanol to avoid the formation of the toxic phosphonothioic acid 1d. Household bleach has also been used routinely in the laboratory to detoxify VX. In both cases, the hypochlorite anion probably acts simultaneously as an oxidant for the sulfur and as a nucleophile attacking the phosphorus.¹¹ However, the reaction stoichiometry has never been fully determined.¹² Indeed, very little published work can be found on the oxidation mechanisms of VX.¹³ Prior work has concentrated on the oxidative chlorination of the agent in both dilute solutions¹⁴ and bulk quantities.¹⁵

This study is the first systematic investigation of the kinetics and mechanisms of the oxidation of phosphonothiolates. It was undertaken in order to facilitate the development of oxidation systems that would be effective in detoxifying VX at ambient temperatures. Both the sulfur and the nitrogen in VX are potential sites for oxidation which could lead to subsequent decomposition and detoxification of the agent. All of the oxidants used in this study were free from chlorine to concentrate solely on the oxidation reactions. Stable commercial peroxygen compounds [m-chloroperoxybenzoic acid, t-butyl hydroperoxide, and OXONE (active component HSO_5^-)]¹⁶ were used in protic solvents ranging from pure t-butanol to t-butanol-water mixtures to pure water. A selective organic oxidant, N-sulfonyloxaziridine, 4a, was also used for comparison with the peroxyacids which are known to be less selective. In all cases, the oxidation of VX was compared with that of a simpler phosphonothiolate derivative, O,S-diethyl methylphosphonothiolate, 1e (^{31}P : δ 63.4, water), so that the effect of the diisopropylamino moiety (or the nitrogen) of VX could be ascertained. It was observed that while a significant amount of 1e could be dissolved in pure water, it remained unchanged for at least 7 months. Therefore, 1e, which is hydrolytically unreactive, makes an excellent substrate for oxidation studies in aqueous solutions.

III. EXPERIMENTAL SECTION.

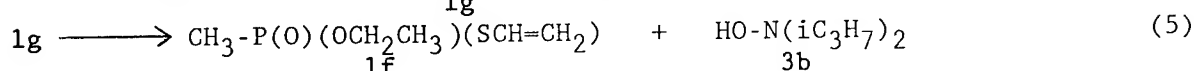
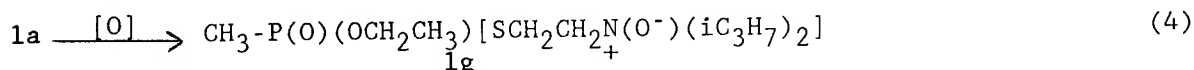
Materials: The phosphonothiolate esters, VX, 1e, and 1h were prepared in-house by Mr. L. J. Szafraniec, Research Directorate, CRDEC and the N-sulfonyloxaziridine, 4a, was obtained from Dr. F. A. Davis, Drexel University. t-Butyl alcohol (Fisher) was freshly distilled and stored over molecular sieve. All other oxidants and solvents were purchased and used as received. The active oxygen equivalent per mole of OXONE, $2KHSO_5 \cdot K_2SO_4 \cdot KHSO_4$, (1.74) and m-chloroperoxybenzoic acid (0.83) was determined by standard titration with sodium thiosulfate.

General Procedure. ^{31}P NMR was used to monitor the oxidations at 19-21°C. For very fast reactions, ^1H NMR was used for substrate concentrations of 0.0005M in D_2O . Typically, a weighed amount of substrate was placed into a 5-mm o.d. Pyrex NMR tube. The appropriate amount of oxidant was weighed into a separate glass vial and dissolved in the appropriate solvent. An aliquot of the oxidizing solution was then pipetted into the NMR tube; the tube was quickly capped, wrapped with Parafilm and shaken to ensure complete mixing of the reactants. The NMR tube was placed in the spectrometer, and spectra were recorded periodically using either a Varian XL-200 or a Varian VXR-400S FTNMR system. All reaction products were identified by ^{31}P and ^{13}C NMR. GC/MS and Direct Exposure Probe (DEP) mass spectrometry were used to assist and confirm the NMR identifications.

IV. RESULTS AND DISCUSSION.

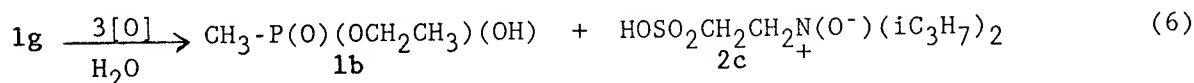
1. Oxidation of the Nitrogen.

In pure t-butanol as well as in a 50 vol% t-butanol solution, an equal molar mixture of VX (0.03M) and m-chloroperoxybenzoic acid reacted instantaneously to form an N-oxide, 1g (^{31}P : δ 59.1, t-butanol/water). (Eq. 4.) One equivalent of active oxygen ($[\text{O}]$) was consumed per mole of VX reacted. As shown in Eq. 5, 1g subsequently decomposed to 1f, O-ethyl S-vinyl methylphosphonothiolate (^{31}P : δ 51.8, t-butanol), and 3b, diisopropyl hydroxylamine, apparently via the cis-elimination of a Cope reaction,^{19,20} with a half-life of ca. 2 hrs at 20°C in pure t-butanol.



As expected for an E1 mechanism, the decomposition half-life of 1g was reduced to about 1 hr in a less-polar solvent system (benzene/t-butanol) while only 10% of 1g decomposed in an equal volume mixture of t-butanol/water after 27 hrs.²¹ This is the first time the N-oxide of VX has been identified! Since the P-S bond is still present in both 1g and 1f, these compounds are believed to be toxic; thus, oxidation at the nitrogen may not completely detoxify VX.

In the presence of excess m-chloroperoxybenzoic acid in 50 vol% t-butanol, the sulfur in 1g was subsequently oxidized followed by immediate hydrolysis at the phosphorus. The sulfinyl group was then oxidized completely to the sulfonic acid, 2c, in Eq. 6. However, ^{13}C NMR indicated that 2c decomposed within a few hours presumably via further oxidation at the nitroso group by the excess oxidant.



The reaction profile shown in Figure 1 clearly demonstrates that the oxidation of the sulfur atom occurred only after all of the VX was converted to the N-oxide, 1g. At 19°C, in 50 vol% t-butanol, the first-order rate constant for sulfur oxidation by excess m-chloroperoxybenzoic acid gave a half-life of 11 min (Table 4). It is believed that the observed reaction was initiated by the oxidation of the sulfur in 1g, followed by a fast hydrolysis to form 1b and 2c; VX was detoxified!

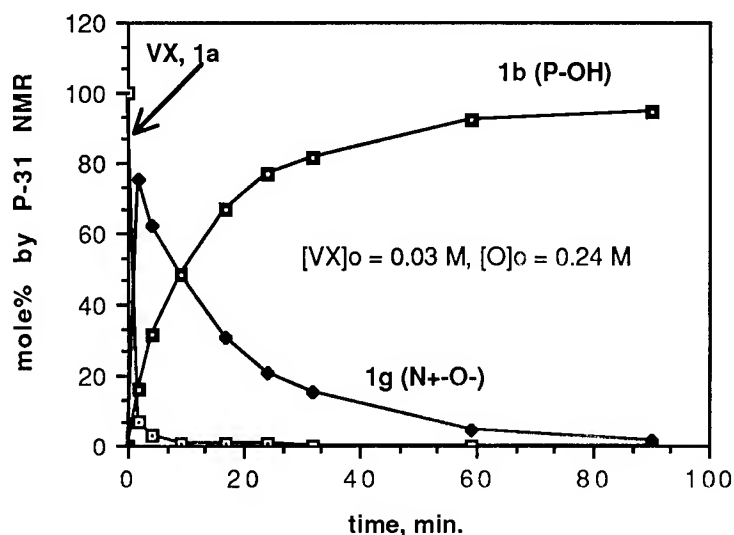


Figure 1. Oxidation of VX by m-Chloroperoxybenzoic Acid in 50 Vol% t-Butanol at 19°C.

Under normal conditions, sulfur should be oxidatively more reactive than nitrogen. However, in VX, the nucleophilicity of the sulfur appears to be reduced probably as a result of the delocalization of electrons to the phosphoryl oxygen. The nitrogen bonded to the 2 isopropyl groups (electron-donating), on the other hand, appears to be more strongly nucleophilic. This hypothesis is supported by the hydrolysis studies of VX reported previously.^{8,9} The diisopropylethyleneimmonium ion (the 3-membered ring mono-cation of 3a in Eq. 2) is believed to be a hydrolysis intermediate which subsequently dimerizes to form 3a. This cyclic immonium ion is formed by the participation of the nitrogen as an internal nucleophile to break the S-C bond. However, the presence of an ethylene-sulfonium ion²² via the participation of the sulfur in the breaking of the C-N bond has never been detected.

2. Oxidation of the Sulfur.

In a mixture of 0.03M 1e and 0.24M m-chloroperoxybenzoic acid in pure t-butanol, the oxidation of the sulfur and the subsequent cleavage

of the P-S bond were further examined. As shown in Figure 2, 1e was first oxidized to a stable intermediate which is believed to be the sulfoxide 1j ($^{31}\text{P}:\delta 26.8$, t-butanol), Eq. 7. Subsequently, 1j reacted to form 3 products: 1k ($^{31}\text{P}:\delta 23.4$, t-butanol), 1l ($^{31}\text{P}:\delta 26.5$, t-butanol) and 1b, Eqs. 8-10. However, the structure of 1l was not positively identified.

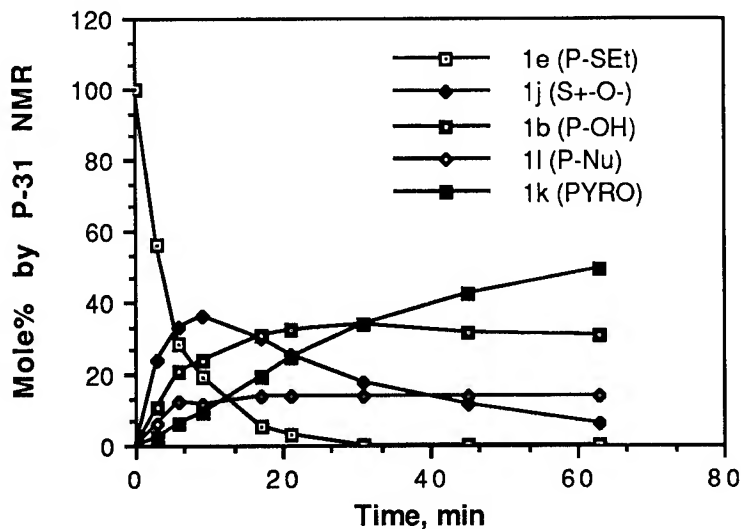
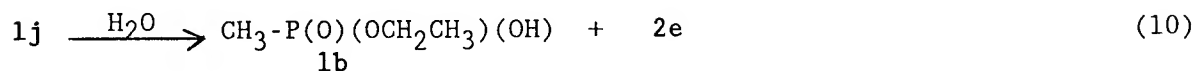
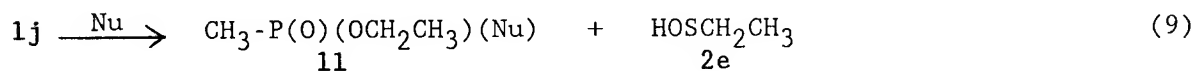
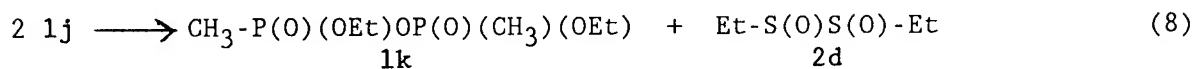
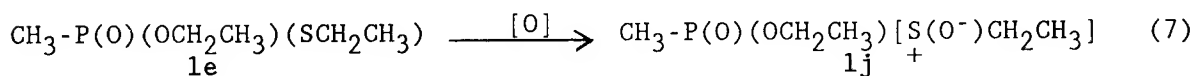
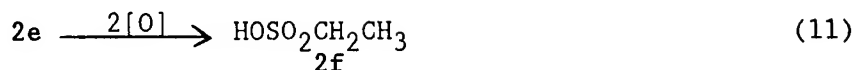


Figure 2. Oxidation of 1e By Excess m-Chloroperoxybenzoic Acid in t-Butanol at 19°C.



As shown in Eq. 9, it may be the product of 1j with any of the following possible nucleophiles (Nu) in the reaction system: m-chlorobenzoate, m-chloroperoxybenzoate, or an impurity present in the oxidant. The ethylsulfenic acid product, 2e, proposed in Eqs. 9-10 was also not detected. Instead, 2f, which formed from the rapid oxidation of 2e, was the major oxidation product (Eq. 11). In fact, competition between 2e and the substrate for the oxidant was observed: A minimum of 3 moles of active oxygen was required to react with 1 mole of 1e indicating that the oxidation of 2e in Eq. 11 is much faster than that of 1e in Eq. 7.



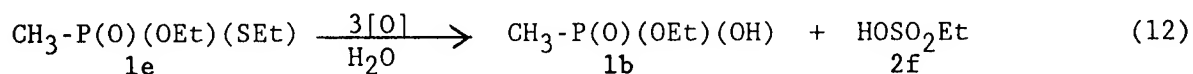
The sulfur in the N-oxide of VX was oxidized in a similar manner. In pure t-butanol, 5 phosphorus-containing compounds were produced from the oxidation of the N-oxide by excess m-chloroperoxybenzoic acid. The major product was 1b since a trace amount of water was apparently present in the solvent. In addition, a significant amount of the pyrophosphate, 1k, was also formed. Of the 3 other compounds, 1 (\$\delta\$26.5) appears to be the sulfoxide of 1g and the other 2 (\$\delta\$28.6 and 39.0) appear to be further displacement products of the sulfoxide which reacts with any nucleophile present in the system.

a. Formation of the Pyrophosphate in the Absence of Water.

The formation of the pyrophosphate, 1k, is important since this toxic compound (LD50, iv rabbit, mg/Kg: 0.089)²³ has often been identified as a major degradation product in neat VX and 1e as well as in solutions of VX and 1e in organic solvents.²⁴ As shown in Figure 2, the rate of production of 1k was initially slow but increased as the concentration of 1b began to decrease. This indicates that 1k was produced only after sufficient amounts of both 1j and 1b were present. The formation of 1k might be initiated by a trace amount of moisture to produce 1b which then acts as a nucleophile to react with another molecule of VX or 1e to produce 1k. Since the toxicity of 1k is significant, a large amount of a nucleophile or water is required to react with 1j in order to achieve detoxification. This subsequent displacement reaction must compete effectively with the formation of 1k.

b. Hydrolysis of the Sulfoxide Intermediate.

In an aqueous solution of 50 vol% t-butanol, 1e reacted quickly with the m-chloroperoxybenzoic acid to produce 1b and ethylsulfonic acid, 2f, in a single step (Eq. 12).



The sulfoxide intermediate, 1j, was too short-lived to be detected by ³¹P NMR; hydrolysis must have occurred as soon as the sulfur was oxidized to form an excellent leaving group. The observed first-order rate constant for Eq. 12 was determined in 3 t-butanol/water mixtures in the presence of 20 times excess m-chloroperoxybenzoic acid. The half-lives were 1.0, 1.6 and 2.1 min, respectively, in t-butanol containing 50, 25 and 10 vol% water. The small change in the rates demonstrates that as long as a sufficient amount of water is present, the subsequent hydrolysis step is

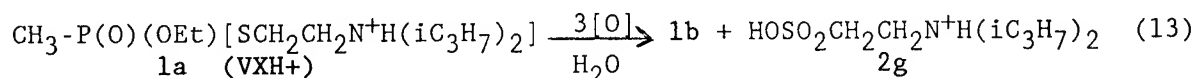
much faster than the oxidation step. Since the sulfoxide intermediate is relatively polar, the measured oxidation rate decreases slightly as the polarity of the solvent decreases. In addition, it should be noted that the sulfur in 1e is more reactive than the sulfur in VX. The oxidation rate of 1e in 50 vol% t-butanol was 5 times faster than that of the sulfur in the N-oxide of VX (Table 4). The identification of 1j as a stable oxidation intermediate, the absence of hydrolytic reactivity for 1e, and the fast oxidation rates for both VX and 1e in aqueous solutions have made it possible to conclude that hydrolysis of both VX and 1e occur only after the sulfur is oxidized. The alternate mechanism that assumes oxidation of the sulfur is preceded by hydrolysis at the P-S bond is very unlikely.

3. Oxidation by Peroxides.

The oxidation of both VX and 1e by t-butyl hydroperoxide (t-BuOOH) was monitored for a 24 hr period using ^{31}P NMR in 3 solvent systems: t-butanol, water, and 50 vol% t-butanol/water. Neither VX nor 1e was oxidized in any of the 3 solvent systems during this period. The only reaction detected was that of VX in water only: VX hydrolyzed via the same parallel reaction paths (Eqs. 1-3) as in pure water but at a slightly faster rate. This small rate enhancement may be attributed to the subsequent oxidation of the sulfide hydrolysis products. However, no reaction could be detected for 1e. In a separate experiment,²⁵ hydrogen peroxide present in large excess was found to be a weakly reactive oxidant for VX. At room temperature in a mixture of 0.02M VX and 0.85M peroxide in 90 vol% 2-propanol, compounds 1g, 1f as well as 1b were detected by ^{31}P NMR after ca. 18 hrs. The weak reactivity of the above oxidants is not surprising since even thioethers can only be oxidized very slowly by excess amounts of peroxides.²⁶ It is therefore concluded that peroxides, in general, are not effective in detoxifying the phosphonothiolates.

4. Oxidation in Acidic Solutions.

As shown in Eq. 13, VX reacted with OXONE quickly in pure water to form the phosphonic acid 1b and sulfonic acid 2g via exclusive P-S bond cleavage. Only 2 moles of OXONE containing 3 equivalents of active oxygen were required to oxidize 1 mole of VX. Neither the N-oxide of VX nor any other intermediates were detected during the course of the reaction. In the presence of 15 moles excess OXONE, ^{13}C NMR showed that 2g remained unchanged and the nitrogen was not oxidized even after 24 hrs. The single oxidation step in Eq. 13 appears to be due to the fact that the nitrogen cannot be oxidized since it is completely protonated in the OXONE solution which has a pH of 2.3 at 20°C.



The first-order rate constant for the reaction of 0.0005M VX and 0.01M OXONE in D_2O at $19^\circ C$ was found to have a half-life of 19 min. This rate is fast enough to detoxify VX at practical concentrations, (e.g., 0.1M) particularly since VX dissolves quickly in the acidic OXONE solution. In the less polar 50 vol% t-butanol solution, the oxidation rate decreased. In addition to the major products 1b and 2g, about 2% 1g was produced. The protonation of the nitrogen is apparently not as complete in 50% t-butanol as in pure water.

To verify the hypothesis that the nitrogen, once protonated, would not be oxidized, a reaction mixture of 0.01M VX, 0.05M HCl and 0.04M m-chloroperoxybenzoic acid in 50 vol% t-butanol was examined. More than 80% of the VX was oxidized directly to 1b and 2g according to Eq. 13 in less than 30 min. About 20% of the VX was not protonated and was oxidized at the nitrogen first. In contrast, as demonstrated in Figure 1 previously, all of the VX was oxidized at the nitrogen in the absence of HCl. Furthermore, the overall reaction rate was faster in the presence of HCl than in the neutral solution. It is possible that the sulfoxide intermediate is more stable when both the phosphoryl and the sulfinyl oxygens are protonated.

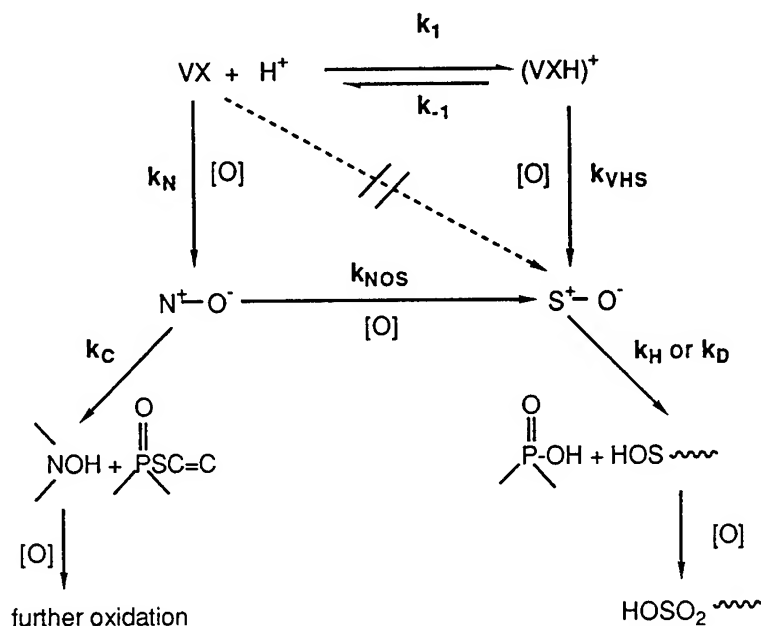
The same products and stoichiometry shown in Eq. 12 above were found for the oxidation of 1e by OXONE in pure water. The oxidation half-life of 0.0005M 1e by 0.01M OXONE in D_2O is 4.8 min, about 4 times faster than that of VX (Table 4). As discussed previously, the m-chloroperoxybenzoic acid also oxidized the sulfur in 1e at a faster rate than the sulfur atom in either the protonated VX or the N-oxide of VX (Table 4). Since 1e does not hydrolyze, the sulfur of 1e must be more tightly bound to the phosphorus relative to the sulfur in VX, which does hydrolyze slowly. One would expect the sulfur in VX to be oxidized faster, but this is contrary to the above observations. The bulky amino group may contribute to the lower oxidative reactivity of the sulfur atom in VX. The greater oxidative reactivity and lack of hydrolytic reactivity of 1e relative to VX again demonstrates that the oxidation of 1e is independent of any prior hydrolysis. In the final acidic reaction mixture in the presence of excess OXONE, 1b can hydrolyze further to methylphosphonic acid ($CH_3P(O)(OH)_2$, 1m) and ethanol via accelerated acid-catalysis.⁸

5. Multiple Paths of VX Oxidation.

Based on the discussions above, the mechanism of VX oxidation by a reactive peroxygen compound is proposed in Scheme 1. In the presence of a reactive oxidant in a neutral organic or neutral aqueous solution, VX is oxidized at the nitrogen first to form an N-oxide (k_N). The N-oxide is stable in aqueous solutions but decomposes in both protic and aprotic organic solvents by a Cope reaction. In aqueous solvents, the sulfur in the N-oxide molecule can be further oxidized by excess oxidant to a sulfoxide intermediate which hydrolyzes at the P-S bond immediately (k_H). The rate of oxidation at the sulfur (k_{NOS}) is smaller than k_N .

In acidic solution, VX is converted to an ammonium salt (VXH^+) and can only be oxidized at the sulfur. The acid oxidation path, k_{VHS} , is faster than the neutral oxidation path, k_{NOS} . In a less polar, acidic solution, on the other hand, a fraction of the VX is unprotonated, and the N-oxide of VX is formed. The equilibrium between VX and VXH^+ (k_1 and k_{-1} ; with pKa of 8.6 at 25°C in pure water)⁸ is affected by the polarity of the solvent system which controls the mechanism and products of VX oxidation.

SCHEME 1



In the 2 extreme cases, organic solvents promote the reactions in the direction of k_{-1} , k_N and k_C (rate of the Cope reaction; negligible in polar solvents); whereas aqueous acidic solutions promote the reactions in

Table 4. Oxidation Half-Lives with Excess Oxidant at 19-20°C.^a

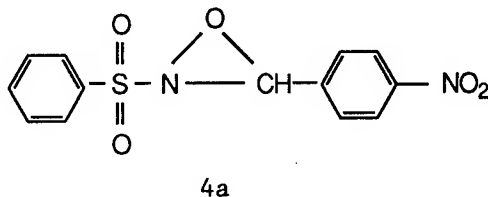
SUBSTRATE	SOLVENT	OXIDANT	[O], M	HALF-LIFE, min
VX	CDCl ₃	4a	0.010	4.1 (k_N)
	50 vol% t-BuOH	peroxybenzoic acid	b	<1 ^b (k_N)
	D ₂ O	OXONE	0.017	19 (k_{VHS})
	50 vol% t-BuOH	peroxybenzoic acid	0.20	11 (k_{NOS})
1e	D ₂ O	OXONE	0.017	4.8 (k_S)
	50 vol% t-BuOH	peroxybenzoic acid	0.44	1 (k_S)

a. The estimated error in the half-lives is $\pm 3-5\%$.

b. The reaction was instantaneous at a stoichiometric concentration of the reactants (0.03M).

the directions of k_1 , k_{VHS} , and k_D (the rate of the displacement reaction at the P-S bond after sulfur is oxidized). In both cases, the fast, secondary oxidation of the sulfide and amine products competes with VX for the oxidant. No direct oxidation of the sulfur in VX was ever observed before the nitrogen was oxidized or protonated. For peroxygen compounds, at least, the rate constant k_g as represented by the dotted arrow in Scheme 1 does not exist. Observed half-lives for each oxidation step in the presence of excess oxidant are summarized in Table 4.

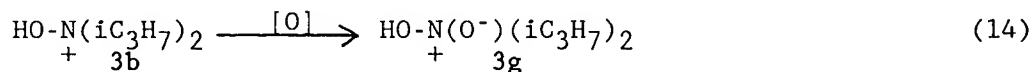
6. Comparison with Oxidation by N-Sulfonyloxaziridine in $CDCl_3$.



In a series of recent publications, n-sulfonyloxaziridines such as 4a were found effective in oxidizing a bivalent sulfur to the sulfoxide in organic solvents.^{17,18} An investigation into the interaction of 4a with the phosphonothiolates was undertaken in order to determine if the sulfur in VX could be oxidized by 4a selectively and if the oxidation mechanism were different from the peroxyacids discussed above. Instead of t-butanol, $CDCl_3$ was used as the solvent because 4a is only sparingly soluble in t-butanol. Similar to the oxidation by the m-chloroperoxybenzoic acid discussed previously, VX was oxidized to 1g which subsequently decomposed to 1f (Eqs. 4-5). The oxidant 4a was reduced to an imine which hydrolyzed slowly to a sulfonamide and a benzaldehyde as a result of a trace amount of water in the solvent. The ^{31}P NMR profile of a reaction mixture containing 2 moles of 4a per mole of VX is illustrated in Figure 3.

a. Secondary Oxidation of the Amine Product.

A different stoichiometry from that of the VX/peroxybenzoic acid oxidation was observed, however. Two moles of 4a were required to react completely with 1 mole of VX. Apparently, as VX was being oxidized, the hydroxylamine was produced at a comparable rate (from the decomposition of 1g) and was able to compete with the VX for the oxidant. Eq. 14.



As discussed previously, only 1 equivalent of the m-chloroperoxybenzoic acid was required per mole of VX in t-butanol because the oxidation was faster and the Cope reaction was slower. Thus, all of the oxidant was

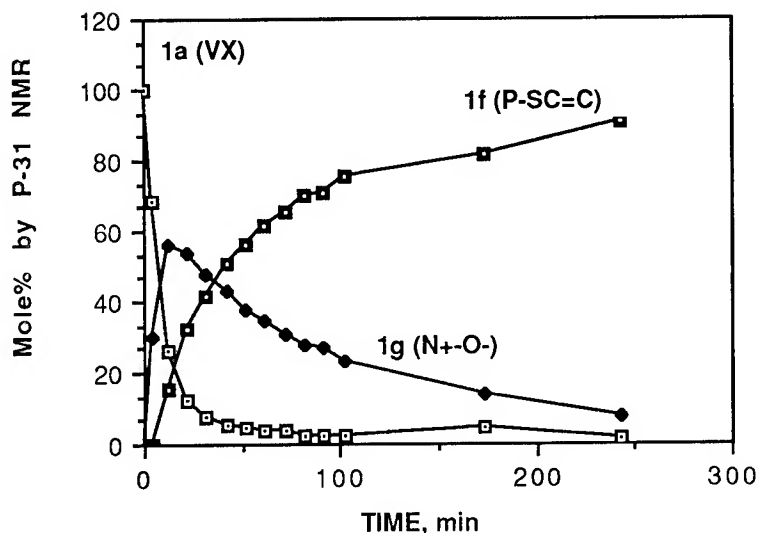
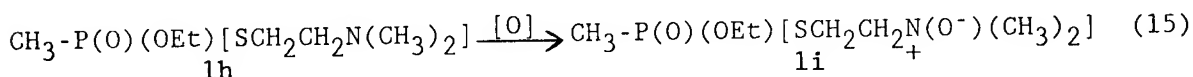


Figure 3. Oxidation of 0.02M VX by 4a in CDCl_3 at 19°C .

consumed by the VX before any of the hydroxylamine could be produced.

The effect of secondary oxidation on reaction stoichiometry was further examined by using a close derivative of VX, O-ethyl S-(dimethylamino)ethyl methylphosphonothiolate, as the substrate (1h, ^{31}P : 856.8, CDCl_3). As shown in Eq. 15, 1h reacted instantaneously with 1 mole of 4a to form 1 mole of the N-oxide (1i, ^{31}P : 856.1, CDCl_3) which was stable in CDCl_3 for at least 24 hrs. No Cope reaction occurred during this period. After 24 hrs, 1i began to hydrolyze slowly and small amounts of both 1b and the pyrophosphate 1k were identified by ^{31}P NMR.



A comparison of the oxidation of 1h with that of VX indicates that both the formation and the decomposition of the N-oxides are sterically controlled. The nitrogen in 1h is expected to be less nucleophilic than that in VX; however, the steric effect appears to predominate. 1i forms at a faster rate but does not decompose by a Cope reaction. For the same reason, a small model compound, diisopropyl methylamine was oxidized instantaneously by 4a to the N-oxide which was stable even under mass spectrometric conditions. Therefore, the hydroxylamine product, 3b, must also oxidize at a faster rate than VX.

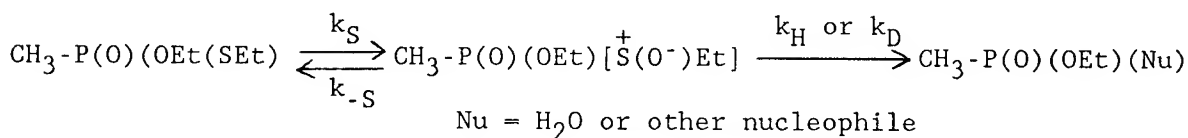
In the presence of 10 moles excess 4a, the half-life for the nitrogen oxidation (k_N) in CDCl_3 at 19°C was measured to be 4.1 min. This rate is slower than the oxidation of the nitrogen by the peroxybenzoic acid by at least 1 order of magnitude (Table 4). Relative to m-chloroperoxybenzoic acid and OXONE, 4a is, therefore, a weaker oxidant. Instead of reacting

as a selective oxidant for the bivalent sulfur, 4a still oxidizes the nitrogen of VX first. This again demonstrates that the nitrogen in VX is more nucleophilic than the sulfur since the reaction mechanism of 4a has been determined to be purely SN_2 .¹⁸ The rate of the decomposition of 1g was independent of the oxidant in $CDCl_3$. Thus, the N-oxide is less stable in the aprotic $CDCl_3$ than in the protic t-butanol solvent in which, as reported above, the decomposition half-life of the N-oxide was ca. 2 hrs. 1f was not further oxidized by excess 4a for at least 24 hrs.

b. Effect of Solvent on Sulfur Oxidation.

1e reacted very slowly with 4a in $CDCl_3$. After 22 hrs, only 36% of 1e was converted to the pyrophosphate 1k. Since the $CDCl_3$ solvent is relatively inert, 1k was found as the only displacement product from the sulfoxide 1j. The slow reaction suggests that 1j is less stable in the less polar organic solvent. Consistent results were obtained from a separate experiment in which 1e was mixed with excess m-chloroperoxybenzoic acid in t-butanol which had been dried over molecular sieve. Different from the reaction profile shown in Figure 2, only small amounts of the 4 products 1b, 1j, 1k and 1l were formed after 2 hrs. Therefore, it is possible to propose a 2-step consecutive reaction mechanism for the sulfur oxidation (Scheme 2). When k_H (hydrolysis) or k_D (displacement) is large, the reverse reaction, k_{-S} , is insignificant in a relatively polar solvent; the rate determining step is the oxidation of the sulfur, k_S . As demonstrated in Section 2, above, when the solvent changed from 50 to 25 to 10 vol% water in t-butanol, the observed first-order rates of oxidation for 1e did not change significantly. The k_H (H_2O) term was apparently larger than k_{-S} in each of these solvent mixtures, and the measured rate constant was the same as k_S . However, in an inert organic solvent such as dry t-butanol or $CDCl_3$ in the absence of any nucleophiles, the k_H (H_2O) or k_D (Nu) term is small and k_{-S} is significant. In this case, both steps in Scheme 2 are rate-determining.

SCHEME 2
Proposed Oxidation Mechanism of 1e



V. SUMMARY.

This paper describes the first systematic attempt to determine the oxidation mechanism of V-type compounds and the controlling factors required for their oxidative detoxification. It was found that the

presence of 2 reactive sites, the strong dependence on the nature of the solvent, and the secondary oxidation of the initial products all contribute to the complexity of the oxidation chemistry of VX. Our findings have resulted in redirecting current research/development involving the chemical decontamination of VX and other organophosphorus nerve agents. Furthermore, the OXONE/water system evaluated in this study is being considered as a replacement for the current ethanolic/HTH method of VX decontamination. Toxicity testing on VX/OXONE solutions is scheduled for early 1990. A package requesting approval to substitute this method and describing the advantages of OXONE over HTH (i.e., increased solubility of VX, a single phase liquid product, no formation of the toxic hydrolysis product, and over 60% less liquid waste when detoxifying 50g or less of VX) is being prepared by CRDEC for submission to the State of Maryland Office of Environmental Programs this fiscal year.

REFERENCES

1. Jackson, J. B., Development of Decontaminating Solution DS2. U.S. Army Chemical Warfare Laboratories Technical Report. CWLR 2368. April 1960. UNCLASSIFIED Report.
2. Day, S. E. and Gendason, P., Evaluation of Air Force Decontaminant CD-1. Chemical Systems Laboratory Technical Report. ARCSL-TR-79054. October 1979. UNCLASSIFIED Report.
3. Research Directorate, CRDEC. To Support the Delisting of Decontaminated Liquid Chemical Surety Materials as Hazardous Waste from Specific Sources Comar 10.51.02.16 & 10.51.02.16-1 K-991-K999 & Residues from K991-K999 February 1988. UNCLASSIFIED Report-Limited Distribution.
4. Title 26 Department of the Environment, Subtitle 13. Disposal of Controlled Hazardous Substances. 26.13.02. Identification and Listing of Hazardous Waste. Maryland Register, Vol. 16, Issue 4, 24 February 1989, p. 498-499.
5. O'Brien, R. D., Toxic Phosphorus Esters. 1960, Academic Press, London, Chapter 2.
6. Jenks, W. P.; Gilchrist, M., J. Am. Chem. Soc., 1964, 86, 5616-5620.
7. DeBruin, K. E.; Tang, C.-I.; Johnson, D. M.; Wilde, R. L., J. Am. Chem. Soc., 1989, 111, 5871-5879.
8. Epstein, J.; Callahan, J. J.; Bauer, V. E., Phosphorus, 1974, 4, 157-163.
9. Szafraniec, L. J.; Szafraniec, L. L.; Beaudry, W. T.; Ward, J. R., On the Stoichiometry of Phosphonothiolate Ester Hydrolysis, in Proceedings of the CRDEC Scientific Conference on Chemical Defense, U.S. Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD, November, 1988.
10. a. For VX, 1a: Edgewood Arsenal Special Report, EO-SR-74001, Chemical Agent Data Sheets, Vol. I, December, 1974. b. For 1d: Michel, H. O.; Epstein, J.; Plappinger, R. R.; Fleisher, J. H., EA 2192 -A Novel

Anticholinesterase, U.S. Army Chemical Research and Development Laboratory Technical Report CRDLR 3135, May 1962, UNCLASSIFIED Report.

11. Epstein, J.; Bauer, V. E.; Saxe, M.; Demek, M. M., J. Am. Chem. Soc., 1956, 78, 4046-4071.

12. a. Horner, L.; Gerhard, J., Phosphorus and Sulfur, 1985, 22, 13-21.
b. Preliminary NMR studies in this laboratory show that VX can only be converted to the phosphonic acid, 1b, by a large excess of the active chlorine from the bleach solution. The non-phosphorus products have not been identified. Sulfoxide, sulfone, ether and carbonyl groups may be present in these products.

13. Poziomek, E. J.; Hackley, B. E., Jr.; Steinberg, G. M., J. Org. Chem., 1958, 23, 714.

14. Bauer, V. E.; Demek, M. M.; Corey, G. G.; Epstein, J., Chlorination of VX in Dilute Aqueous Solution at 25°C, U.S. Army Chemical Warfare Laboratory Technical Report CWLR 2231. August 1958. UNCLASSIFIED Report.

15. Benson, J. H.; Valis, R. J.; Kramer, S. R.; Heminger, D. M.; Hildebrandt, H. F.; Berenthien, T. C., Laboratory and Pilot Scale Detoxification of VX in Acidic Media Using Chlorine Gas, Edgewood Arsenal Technical Report EM-TR-73070, May 1974. UNCLASSIFIED Report.

16. For a review: Kennedy, R. J.; Stock, A. M., J. Org. Chem., 1960, 25, 1901-1907.

17. David, F. A.; Jenkins, Jr., R. H.; Yocklovich, S. F., Tet. Lett., 1978, 5171-5174.

18. David, F. A.; Billmers, J. M.; Gosciniak, D. J.; Towson, J. C.; Bach, R. D., J. Org. Chem., 1986, 51, 4240-4245.

19. Cope, A. C.; Pike, R. A.; Spencer, C. F., J. Am. Chem. Soc., 1953, 75, 3212.

20. Cope, A. C.; Lebel, N. A.; Lee, H. H.; Moore, W. R., J. Am. Chem. Soc., 1957, 79, 4720.

21. After 27 hrs, 1g also began to hydrolyze in the 50 vol% t-butanol solution to product 1b which was identified by ³¹P NMR. The hydrolysis rate was slightly faster than VX indicating that the oxidation at the nitrogen makes the sulfur moiety of VX a slightly better leaving group.

22. Yang, Y.-C.; Szafraniec, L. L.; Beaudry, W. T.; Ward, J. R., J. Org. Chem., 1988, 53, 3293-3297.

23. Private communication: Mr. Leonard J. Szafraniec, Research Directorate, CRDEC; unpublished data.

24. D'Agostino, P. A.; Provost, L. R.; Visentini, J., J. Chromatogr., 1987, 402, 221-232.

25. Albizo, J. M. and Szafraniec, L. L., CRDEC, unpublished data.

26. Curci, R. and Edwards, J. O., Organic Peroxides (Swern, D., Ed.), Wiley-Interscience, New York, N.Y., 1970, 1, 230-240.